

Large Deviations for Brownian Particle Systems with Killing.*

Amarjit Budhiraja and Wai-Tong (Louis) Fan and Ruoyu Wu

April 18, 2017

Abstract: Particle approximations for certain nonlinear and nonlocal reaction-diffusion equations are studied using a system of Brownian motions with killing. The system is described by a collection of i.i.d. Brownian particles where each particle is killed independently at a rate determined by the empirical sub-probability measure of the states of the particles alive. A large deviation principle (LDP) for such sub-probability measure-valued processes is established. Along the way a convenient variational representation, which is of independent interest, for expectations of nonnegative functionals of Brownian motions together with an i.i.d. sequence of random variables is established. Proof of the LDP relies on this variational representation and weak convergence arguments.

AMS 2000 subject classifications: Primary 60F10, 60K35; secondary 60B10, 93E20.

Keywords: Large Deviations, Weakly interacting particle systems, Brownian particles with killing, nonlinear reaction-diffusion equations, variational representations.

1. Introduction

The goal of this work is to study a large deviation principle (LDP) for an interacting particle system associated with the nonlinear and nonlocal reaction-diffusion equation

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2} \Delta u(t, x) - \langle \zeta, u \rangle u(t, x), \quad (t, x) \in (0, \infty) \times \mathbb{R}^d, \quad \lim_{t \downarrow 0} u(t, \cdot) = \delta_0(\cdot), \quad (1.1)$$

where Δ is the d -dimensional Laplacian operator, δ_0 is the Dirac measure at 0 and $\zeta: \mathbb{R}^d \rightarrow \mathbb{R}_+$ is a continuous function with sub-quadratic growth, namely

$$0 \leq \zeta(x) \leq C_\zeta(1 + \|x\|^p) \text{ for all } x \in \mathbb{R}^d \quad (1.2)$$

for some $0 \leq p < 2$ and $C_\zeta \in (0, \infty)$. Here $\langle \zeta, u \rangle$ denotes the integral $\int_{\mathbb{R}^d} \zeta(x) u(t, x) dx$. Roughly speaking the system is described by n independent Brownian particles where each particle is killed independently at a rate determined by the empirical measure of current particle states. More precisely, let $\{X_i\}_{i \geq 1}$ be a sequence of i.i.d. exponential random variables with rate 1 and let $\{B_i(t), t \geq 0\}_{i \geq 1}$ be independent d -dimensional standard Brownian motions independent

*Research supported in part by the National Science Foundation (DMS-1305120), the Army Research Office (W911NF-14-1-0331) and Defense Advanced Research Projects Agency (W911NF-15-2-0122).

of $\{X_i\}_{i \geq 1}$. Define for $t \geq 0$ the random sub-probability measure $\mu^n(t)$ as the solution to the following equation

$$\mu^n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{B_i(t)} \mathbf{1}_{\{X_i > \int_0^t \langle \zeta, \mu^n(s) \rangle ds\}}. \quad (1.3)$$

Since a.s., we can enumerate $\{X_i\}_{i=1}^n$ in a strictly increasing order, the unique solution of (1.3) can be written explicitly in a recursive manner. It can be checked (see Theorem 2.1) that $\mu^n \doteq \{\mu^n(t)\}_{t \in [0, T]}$ converges, in the Skorokhod path space \mathcal{D} , in probability to μ where for $t > 0$, $\mu(t)$ has density $u(t, \cdot)$ given as the solution of (1.1). Such particle systems are motivated by problems in biology, ecology, chemical kinetics, etc. For example, the simplest case where $\zeta \equiv 1$ corresponds to the case where the killing rate is proportional to the total number of particles alive and models a setting in which particles compete for a common resource. More general functions ζ are of interest as well and one interpretation of $\zeta(x)$ is the amount of resource consumed by a particle in state x (see for instance [16] and Chapter 9 of [24]). Similar particle systems arise in problems of mathematical finance as models for self exciting correlated defaults [6].

In this work we are interested in a Large Deviation Principle for the stochastic process μ^n in \mathcal{D} . A LDP, in addition to giving precise estimates on exponential decay rates of probabilities of deviations of the stochastic system from its law of large number limit, is useful for developing accelerated Monte-Carlo methods for rare event simulation (see eg. [12, 11]). Large deviation results for particle systems associated with nonlinear reaction-diffusion equations have been studied in many works [17, 1, 20]. However, these settings are very different from ours. Specifically, in these works the models are described through jump-Markov processes on a discrete lattice with local interaction rules, the reaction operators in the PDE are local, and proof ideas are quite different. The setting considered in the current work is closer to that of particle systems with a weak interaction. Large deviation results for weakly interacting diffusions go back to the classical work of Dawson and Gärtner [7] (see also Chapter 13 of [15]) who studied uniformly nondegenerate diffusions that interact through a drift term that depends on the empirical measure of the particle states. In a recent work [4], using certain variational representations from [2] for exponential functionals of Brownian motions, a different proof of the LDP in [7] was provided. The work in [4] in particular relaxed the assumption of uniform nondegeneracy and allowed mean field interaction in both drift and diffusion coefficients. Unlike [7], proofs in [4] do not involve any space-time discretization or exponential probability estimates but rather rely on certain stochastic control representations and weak convergence arguments. A similar approach is taken in the current work although the interaction here is of a different form. The representation in [2] is not well suited for our setting, since our particle system depends not only on the Brownian motions $\{B_i\}$ but also on the i.i.d. sequence $\{X_i\}$. To address this, we begin by establishing a new variational representation (Proposition 4.1) that allows the functional to depend, in addition to the Brownian motions, on an extra collection of i.i.d. random variables. This representation plays a central role in the proof of the LDP and we believe that it is of independent interest and has potential applications for large deviation analysis of other interacting particle models (see Remark 3.2 for one such application). The starting point for the proof of the representation is the Boué-Dupuis [2] variational formula and the classical Donsker-Varadhan representation for exponential integrals. The key step in the proof is the interchange of a certain ‘integration’ and ‘infimum’ operation (see discussion

above Proposition 4.1) for which we need a careful measurable selection and construction of suitable controls (see Section 9). Using the Laplace formulation of the LDP (see Chapter 1 of [10]) the variational representation in Proposition 4.1 reduces the proof of the upper bound in the LDP to proving tightness and characterization of limit points of certain controlled empirical measures. The proof of the lower bound relies on construction of asymptotically near optimal controlled stochastic processes with killing, together with a uniqueness argument for limit point that crucially uses the form of the killing function (see Lemma 3.1).

Although not explored here, we expect that similar techniques will be useful for settings where particle dynamics are given through general jump-diffusions and the killing function ζ is state dependent, i.e. the integral on the right side of (1.3) is replaced with $\int_0^t \langle \zeta_{B_i(s)}, \mu^n(s) \rangle ds$ where $\zeta(x, y) = \zeta_x(y)$ is a suitable \mathbb{R}_+ valued function on $\mathbb{R}^d \times \mathbb{R}^d$. We note that μ^n can be written as a measurable function of the empirical measure of an i.i.d. collection. More precisely, denoting $\Gamma^n \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(B_i, X_i)}$, we can represent $\mu^n = h(\Gamma^n)$ for a measurable map from $\mathcal{P}(\mathcal{C}([0, T] : \mathbb{R}^d) \times [0, \infty))$ to $\mathbb{D}([0, T] : \mathcal{M}(\mathbb{R}^d))$ (see notation in Section 2). Sanov's theorem gives a LDP for Γ^n , however the function h is not continuous due to which one cannot apply the contraction principle to deduce a LDP for μ^n . Another approach that is common in the study of large deviations for weakly interacting particle systems is to consider a convenient absolutely continuous change of measure under which the large deviation analysis is more tractable and then deduce the LDP for the original collection by obtaining suitable estimates on the Radon-Nikodym derivative. In the current setting one way to implement such an approach would be to first appeal to Sanov's theorem for a LDP for $\tilde{\Gamma}^n \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(B_i, \tilde{\sigma}_i)}$, where $\tilde{\sigma}_i \doteq \inf\{t : X_i \leq \int_0^t \langle \zeta, \mu(s) \rangle ds\}$ for $i = 1, \dots, n$ and μ as before is the law of large number limit of μ^n , and then establish a LDP for $\bar{\Gamma}^n \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(B_i, \bar{\sigma}_i^n)}$, where $\bar{\sigma}_i^n \doteq \inf\{t : X_i \leq \int_0^t \langle \zeta, \mu^n(s) \rangle ds\}$ for $i = 1, \dots, n$, by estimating the Radon-Nikodym derivative $\frac{dQ^n}{dP^n}$ where Q^n and P^n are the probability laws of $\tilde{\Gamma}^n$ and $\bar{\Gamma}^n$ respectively. However obtaining the required estimates on $\frac{dQ^n}{dP^n}$ for carrying out this approach does not appear to be straightforward.

Finally, we remark that both the LDP result and the law of large numbers result are with respect to an “annealed measure” \mathbf{P} (defined in Section 2) under which the randomness of $\{X_i\}_{i \in \mathbb{N}}$ is averaged out. Although not pursued here, it is also of interest to consider a “quenched” version of the LDP, in analogy to results for random walks in random environments and for random polymer models; see for instance [5] and [13] respectively.

The paper is organized as follows. Section 2 introduces the basic model and notation. In this section we also present a law of large numbers (Theorem 2.1) for μ^n . Section 3 gives the main large deviations result of this work, namely Theorem 3.2. Rest of the work is devoted to the proof of Theorem 3.2. We begin in Section 4 by presenting a variational representation for functionals of independent Brownian motions and an i.i.d. sequence of random variables. Proof of this representation is given in Section 9. Tightness of controls and related processes in this representation is argued in Section 5. Using the representation and the tightness results from Section 5, Section 6 proves the Laplace principle upper bound for the LDP in Theorem 3.2 while Section 7 gives the proof of the lower bound. Compactness of sub-level sets of the candidate rate function is established in Section 8. The large deviation principle for μ^n is immediate on combining results of Sections 6, 7 and 8. Finally Appendix A provides the proof of Theorem 2.1.

2. Model and notation

The following notation will be used. For a Polish space $(\mathbb{S}, d(\cdot, \cdot))$, denote the corresponding Borel σ -field by $\mathcal{B}(\mathbb{S})$. For a signed measure ν on \mathbb{S} and ν -integrable function $f: \mathbb{S} \rightarrow \mathbb{R}^d$, we will use $\langle \nu, f \rangle$ and $\langle f, \nu \rangle$ interchangeably for $\int f d\nu$. Let $\mathcal{P}(\mathbb{S})$ [resp. $\mathcal{M}(\mathbb{S})$] be the space of probability measures [resp. sub-probability measures] on \mathbb{S} , equipped with the topology of weak convergence. A convenient metric for this topology is the bounded-Lipschitz metric d_{BL} , which metrizes $\mathcal{P}(\mathbb{S})$ [resp. $\mathcal{M}(\mathbb{S})$] as a Polish space, defined as

$$d_{BL}(\nu_1, \nu_2) \doteq \sup_{\|f\|_{BL} \leq 1} |\langle \nu_1 - \nu_2, f \rangle|, \quad \nu_1, \nu_2 \in \mathcal{P}(\mathbb{S}) \text{ [resp. } \mathcal{M}(\mathbb{S})].$$

Here $\|\cdot\|_{BL}$ is the bounded Lipschitz norm, i.e. for $f: \mathbb{S} \rightarrow \mathbb{R}$,

$$\|f\|_{BL} \doteq \max\{\|f\|_\infty, \|f\|_L\}, \quad \|f\|_\infty \doteq \sup_{x \in \mathbb{S}} |f(x)|, \quad \|f\|_L \doteq \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

For $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{S})$, we write $\nu_1 \ll \nu_2$ if ν_1 is absolutely continuous with respect to ν_2 . Denote by $\mathbb{M}_b(\mathbb{S})$ [resp. $\mathbb{C}_b(\mathbb{S})$] the space of real bounded $\mathcal{B}(\mathbb{S})/\mathcal{B}(\mathbb{R})$ -measurable functions [resp. real bounded and continuous functions]. Let $\mathbb{C}_b^k(\mathbb{R}^d)$ be the space of functions on \mathbb{R}^d , which have continuous and bounded derivatives up to the k -th order.

Fix $T < \infty$. All stochastic processes will be considered over the time horizon $[0, T]$. We will use the notations $\{X_t\}$ and $\{X(t)\}$ interchangeably for stochastic processes. For a Polish space \mathbb{S} , denote by $\mathbb{C}([0, T] : \mathbb{S})$ [resp. $\mathbb{D}([0, T] : \mathbb{S})$] the space of continuous functions [resp. right continuous functions with left limits] from $[0, T]$ to \mathbb{S} , endowed with the uniform [resp. Skorokhod] topology. Probability law of an \mathbb{S} -valued random variable X will be denoted as $\mathcal{L}(X)$. We will use $\mathbf{E}^{\mathbf{P}}$ for expected value under some probability law \mathbf{P} but when clear from the context \mathbf{P} will be suppressed from the notation. We say a collection $\{X_n\}$ of \mathbb{S} -valued random variables is tight if $\{\mathcal{L}(X_n)\}$ are tight in $\mathcal{P}(\mathbb{S})$. Convergence of a sequence $\{X_n\}$ of \mathbb{S} -valued random variables in distribution to X will be written as $X_n \Rightarrow X$.

We will usually denote by $\kappa, \kappa_1, \kappa_2, \dots$, the constants that appear in various estimates within a proof. The value of these constants may change from one proof to another. Let $\mathcal{C} \doteq \mathbb{C}([0, T] : \mathbb{R}^d)$, $\mathcal{S} \doteq \mathcal{C} \times \mathbb{R}_+$ and $\mathcal{D} \doteq \mathbb{D}([0, T] : \mathcal{M}(\mathbb{R}^d))$. Denote by \mathbb{N}_0 the set of nonnegative integers.

Let $\{B_i\}_{i \in \mathbb{N}}$ be a sequence of i.i.d. d -dimensional standard Brownian motions, and $\{X_i\}_{i \in \mathbb{N}}$ be a sequence of i.i.d. mean 1 exponential random variables independent of $\{B_i\}_{i \in \mathbb{N}}$, defined on $(\Omega, \mathcal{F}, \mathbf{P})$. Let \mathcal{F}_t be the \mathbf{P} -completion of the σ -field generated by $\{B_i(s), X_i, s \leq t, i \in \mathbb{N}\}$. Fix a continuous function $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}_+$ satisfying (1.2). The main result of this work is a large deviation principle (LDP) for $\mu^n \doteq \{\mu^n(t)\}_{t \in [0, T]}$ in \mathcal{D} , where $\mu^n(t)$ is defined in (1.3). We begin with the following law of large numbers result.

Let $\theta \doteq \mathcal{L}(X_1)$ be the exponential distribution (with rate 1). Let $\mu_0 \doteq \mathcal{L}(B_1)$ be the Wiener measure on \mathcal{C} , $\mu_{0,t} \doteq \mathcal{L}(B_1(t))$ be the marginal distribution of μ_0 at time instant t , and $b(t) \doteq \langle \zeta, \mu_{0,t} \rangle$. Let $a(t)$ be the unique solution of the ordinary differential equation (ODE)

$$\dot{a}(t) = -a^2(t)b(t), \quad a(0) = 1, \tag{2.1}$$

namely $a(t) = \frac{1}{1 + \int_0^t b(s) ds}$. Note that from (1.2) it follows that

$$\zeta(x) \leq \tilde{C}_\zeta(1 + \|x\|^2), \quad x \in \mathbb{R}^d \quad (2.2)$$

for some $\tilde{C}_\zeta \in (0, \infty)$. In particular $b(t)$ and $a(t)$ are well-defined. Define $\mu: [0, T] \rightarrow \mathcal{M}(\mathbb{R}^d)$ as $\mu(t) \doteq a(t)\mu_{0,t}$. Clearly, $\mu \in \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$ and for positive t , $\mu(t)$ is absolutely continuous with respect to Lebesgue measure. Writing for $t \in (0, T]$, $\mu_t(dx) = u(t, x)dx$, it is easily checked that $u(t, x)$ is the solution to (1.1). Furthermore we have the following law of large numbers whose proof is provided in Appendix A.

Theorem 2.1. *As $n \rightarrow \infty$, μ^n converges to μ in probability in \mathcal{D} .*

By a more careful analysis, for the proof of Theorem 2.1, the assumption that ζ has sub-quadratic growth can be weakened to allow for quadratic growth. However in the proof of the large deviation result that we present in the next section, the condition on sub-quadratic growth of ζ is used in an important way (see proof of Lemma 5.5) and thus we make this assumption throughout.

3. Large deviation principle

In this section we present our main large deviations result. We begin by introducing some canonical spaces and processes that will be convenient in the analysis.

3.1. Canonical space and processes

Let \mathcal{R}^W be the space of finite Borel measures r on $\mathbb{R}^d \times [0, T]$ such that $r(\mathbb{R}^d \times [0, t]) = t$ for all $t \in [0, T]$ and

$$\int_{\mathbb{R}^d \times [0, T]} \|y\| r(dy dt) < \infty.$$

Such a measure can be disintegrated as $r(dy dt) = r_t(dy) dt$, where $t \mapsto r_t(\cdot)$ is a measurable map from $[0, T]$ to $\mathcal{P}(\mathbb{R}^d)$. We equip \mathcal{R}^W with the topology of weak convergence plus convergence of first moments. This topology can be metrized with the Wasserstein distance of order 1 and the space with this metric is Polish (cf. [21, Section 6.3]).

We now introduce the *canonical space* $\Xi \doteq \mathcal{C} \times \mathcal{C} \times \mathcal{R}^W \times \mathbb{R}_+$ and *canonical variables* on $(\Xi, \mathcal{B}(\Xi))$: For $\xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \in \Xi$,

$$\tilde{b}(\xi) \doteq \tilde{b}, \quad b(\xi) \doteq b, \quad \rho(\xi) \doteq \rho, \quad \tilde{\sigma}(\xi) \doteq \tilde{\sigma}. \quad (3.1)$$

Let $\{\mathcal{G}_t\}_{0 \leq t \leq T}$ be the canonical filtration on $(\Xi, \mathcal{B}(\Xi))$, namely

$$\mathcal{G}_t \doteq \sigma\{\tilde{b}_s, b_s, \rho(A \times [0, s]), s \leq t, A \in \mathcal{B}(\mathbb{R}^d), \tilde{\sigma}\}, \quad t \in [0, T].$$

For $\Theta \in \mathcal{P}(\Xi)$, denote by $\Theta_{(i)}$ the marginal of Θ on the i -th coordinate, $i = 1, 2, 3, 4$. For example $\Theta_{(1)}(A) \doteq \Theta(A \times \mathcal{C} \times \mathcal{R}^W \times \mathbb{R}_+)$ for all $A \in \mathcal{B}(\mathcal{C})$.

3.2. Rate function and statement of the LDP

Let \mathcal{P}_∞ denote the collection of all $\Theta \in \mathcal{P}(\Xi)$ such that $\Theta_{(4)} \ll \theta$ and under Θ the following hold:

(1) \mathbf{b} is a d -dimensional standard \mathcal{G}_t -Brownian motion.

(2) $\int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy dt) < \infty$, a.s.

(3) $\tilde{\mathbf{b}}_t = \mathbf{b}_t + \int_{\mathbb{R}^d \times [0, t]} y \boldsymbol{\rho}(dy ds)$, for all $t \in [0, T]$, a.s.

Remark 3.1. Note that the property $\Theta_{(4)} \ll \theta$ above implies that for $\Theta \in \mathcal{P}_\infty$, $\Theta_{(4)}$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}_+ . This observation will be used in the proof of Lemmas 3.1, 5.4 and 5.5.

For $\Theta \in \mathcal{P}(\Xi)$ let

$$\mathcal{J}(\Theta) \doteq \mathbf{E}^\Theta \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy dt) \right] + R(\Theta_{(4)} \|\theta),$$

where $R(\gamma \|\theta)$ for $\gamma \in \mathcal{P}(\mathbb{R}^d)$ is the relative entropy of γ with respect to θ defined as

$$R(\gamma \|\theta) \doteq \begin{cases} \int_{\mathbb{R}^d} \left(\log \frac{d\gamma}{d\theta} \right) d\gamma, & \gamma \ll \theta, \\ \infty, & \text{otherwise.} \end{cases}$$

Note that with Θ_0 the probability law of $(B_1, B_1, \delta_0(dy) \otimes dt, X_1)$, where B_1 is a standard d -dimensional Brownian motion and X_1 is a mean one exponential random variable independent of B_1 , $\mathcal{J}(\Theta_0) = 0$.

For $\Theta \in \mathcal{P}_\infty$ consider the equation, whose unknown is $\varpi_\Theta: [0, T] \rightarrow \mathcal{M}(\mathbb{R}^d)$,

$$\langle f, \varpi_\Theta(t) \rangle = \mathbf{E}^\Theta \left[f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \varpi_\Theta(s) \rangle ds\}} \right], \text{ for all } f \in \mathbb{C}_b(\mathbb{R}^d), t \in [0, T]. \quad (3.2)$$

Note that if a $\varpi_\Theta: [0, T] \rightarrow \mathcal{M}(\mathbb{R}^d)$ satisfies the above equation for all continuous and bounded f then it also satisfies the equality for all bounded measurable f since given a bounded function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ we can find functions $f_n \in \mathbb{C}_b(\mathbb{R}^d)$ such that $f_n(x) \rightarrow f(x)$ for a.e. $x \in \mathbb{R}^d$ and $\{f_n\}$ is uniformly bounded. By monotone convergence theorem we then see that the above equality holds for all non-negative measurable functions from \mathbb{R}^d to \mathbb{R} and finally the equality also holds for all measurable $f: \mathbb{R}^d \rightarrow \mathbb{R}$ that satisfy $\mathbf{E}^\Theta |f(\tilde{\mathbf{b}}_t)| < \infty$.

The following lemma shows that if $\mathcal{J}(\Theta) < \infty$ there is at most one solution to (3.2).

Lemma 3.1. *Suppose $\Theta \in \mathcal{P}_\infty$ satisfies $\mathcal{J}(\Theta) < \infty$. Then there is at most one measurable map $\varpi_\Theta: [0, T] \rightarrow \mathcal{M}(\mathbb{R}^d)$ that satisfies (3.2). Moreover such a solution must necessarily be in $\mathcal{C}([0, T]: \mathcal{M}(\mathbb{R}^d))$.*

Proof. Let $\varpi_\Theta: [0, T] \rightarrow \mathcal{M}(\mathbb{R}^d)$ be a measurable map that satisfies (3.2). Fix $t \in [0, T]$. From

(2.2) and properties of \mathcal{P}_∞ ,

$$\begin{aligned} \mathbf{E}^\Theta \left[\zeta(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \varpi_\Theta(s) \rangle ds\}} \right] &\leq \tilde{C}_\zeta \mathbf{E}^\Theta(1 + \|\tilde{\mathbf{b}}_t\|^2) \\ &\leq 2\tilde{C}_\zeta \left[1 + \mathbf{E}^\Theta \|\mathbf{b}_t\|^2 + \mathbf{E}^\Theta \left\| \int_{\mathbb{R}^d \times [0, t]} y \boldsymbol{\rho}(dy ds) \right\|^2 \right] \\ &\leq \kappa(1 + \mathcal{J}(\Theta)) < \infty. \end{aligned} \quad (3.3)$$

Thus from the comments below (3.2), $\langle \zeta, \varpi_\Theta(t) \rangle$ is finite and equals

$$\mathbf{E}^\Theta \left[\zeta(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \varpi_\Theta(s) \rangle ds\}} \right]$$

for all $t \in [0, T]$. Furthermore, it follows from Remark 3.1 that for each $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t, s \in [0, T]$, as $t \rightarrow s$,

$$f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \varpi_\Theta(u) \rangle du\}} \rightarrow f(\tilde{\mathbf{b}}_s) \mathbf{1}_{\{\tilde{\sigma} > \int_0^s \langle \zeta, \varpi_\Theta(u) \rangle du\}}, \quad \text{a.s. } \Theta,$$

and hence $\langle f, \varpi_\Theta(t) \rangle \rightarrow \langle f, \varpi_\Theta(s) \rangle$. This implies that $\varpi_\Theta \in \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. Now it remains to argue the uniqueness of ϖ_Θ . Let $\pi_1, \pi_2 \in \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$ be two solutions of (3.2). Consider functions $h(t) \doteq \langle \zeta, \pi_1(t) \rangle - \langle \zeta, \pi_2(t) \rangle$ and $H(t) \doteq \int_0^t h(s) ds$ for $t \in [0, T]$. We will argue via contradiction that $H(t) \equiv 0$. This will show $\pi_1 = \pi_2$, proving the desired uniqueness. Suppose $M \doteq \sup_{0 \leq t \leq T} |H(t)| > 0$. Without loss of generality we assume that $M = \sup_{0 \leq t \leq T} H(t)$. Since H is a continuous function, there exists some $t \in [0, T]$ such that $H(t) = M$. Define $t^* \doteq \inf\{t : H(t) = M\}$. Then $H(t^*) = M$. Since $H(0) = 0$, we have $t^* > 0$. Again by continuity of H , there exists $t_0 \in (0, t^*)$ such that for all $t \in [t_0, t^*]$, $H(t) > 0$, and hence

$$\int_0^t \langle \zeta, \pi_1(s) \rangle ds > \int_0^t \langle \zeta, \pi_2(s) \rangle ds, \quad \forall t \in [t_0, t^*].$$

Recalling that $\zeta \geq 0$, we have for all $t \in [t_0, t^*]$,

$$\begin{aligned} h(t) &= \langle \zeta, \pi_1(t) \rangle - \langle \zeta, \pi_2(t) \rangle \\ &= \mathbf{E}^\Theta \left[\zeta(\tilde{\mathbf{b}}_t) \left(\mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi_1(s) \rangle ds\}} - \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi_2(s) \rangle ds\}} \right) \right] \leq 0. \end{aligned}$$

This implies that

$$M = H(t^*) = H(t_0) + \int_{t_0}^{t^*} h(s) ds \leq H(t_0),$$

which contradicts with the definition of t^* . Hence we must have $H(t) \equiv 0$ and this completes the proof. \square

We now introduce the rate function associated with $\{\mu^n\}$. Let \mathcal{P}_∞^1 be the collection of all $\Theta \in \mathcal{P}_\infty$ with $\mathcal{J}(\Theta) < \infty$ and such that there is a solution ϖ_Θ of (3.2) (which is necessarily unique). Define function $I : \mathcal{D} \rightarrow [0, \infty]$ as

$$I(\pi) \doteq \inf_{\{\Theta \in \mathcal{P}_\infty^1 : \varpi_\Theta = \pi\}} \mathcal{J}(\Theta), \quad \pi \in \mathcal{D}. \quad (3.4)$$

As is usual, infimum over an empty set is taken to be ∞ .

The following is the main result of this work.

Theorem 3.2. *The sequence $\{\mu^n\}_{n \in \mathbb{N}}$ satisfies a LDP in \mathcal{D} with rate function I .*

Remark 3.2. The representation in Proposition 4.1 that is key to the proof of Theorem 3.2 can be used to also cover a more general case where the Brownian motions B_i are replaced with $Z_i(\cdot) \doteq Y_i + B_i(\cdot)$ where $\{Y_i\}$ are i.i.d. and independent of $\{B_i, X_i\}$, and the distribution m of Y_1 satisfies suitable integrability conditions. In this case the law of large numbers limit will solve the PDE (1.1) with initial condition m and in the rate function governing the LDP, the term $R(\cdot \parallel \theta)$ will be replaced with a term of the form $R(\cdot \parallel \theta \otimes m)$.

It follows from the second statement of Lemma 3.1 that $I(\pi) = \infty$ for all $\pi \in \mathcal{D} \setminus \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. Hence the effective domain of the rate function is contained in $\mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. By contraction principle (Remark (c) of Theorem 4.2.1 of [8]), one obtains from Theorem 3.2 the LDP in $\mathcal{M}(\mathbb{R}^d)$ for $\{\mu_t^n\}_{n \in \mathbb{N}}$ for each $t \in [0, T]$.

Recalling the equivalence between a LDP and a Laplace principle (cf. [10, Chapter 1]), in order to prove Theorem 3.2 it suffices to show that:

- (1) I defined in (3.4) is a rate function.
- (2) For every $F \in \mathbb{C}_b(\mathcal{D})$,

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))] = \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\}. \quad (3.5)$$

Proof of item (1) is given in Section 8 while the proof of item (2) is carried out in two steps. First in Section 6 we will prove the *Laplace upper bound*:

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))] \geq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\}. \quad (3.6)$$

The proof of (3.5) is then completed in Section 7 by proving the complementary *Laplace lower bound*:

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))] \leq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\}. \quad (3.7)$$

4. A general variational representation formula

In order to prove Theorem 3.2 we need to study the asymptotics of

$$-\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))], \quad (4.1)$$

where $F \in \mathbb{C}_b(\mathcal{D})$. For this we will use certain variational representations. Note that $F(\mu^n)$ can be written as $\Psi(\mathbf{B}^n, \mathbf{X}^n)$, where $\mathbf{B}^n \doteq \{B_i\}_{i=1}^n$ is an nd -dimensional Brownian motion, $\mathbf{X}^n \doteq \{X_i\}_{i=1}^n$ is an \mathbb{R}_+^n -valued random variable and Ψ is a suitable map. When Ψ is just a function of the Brownian motion \mathbf{B}^n , a variational representation for quantities as in (4.1) was obtained in [2] (see also [3] where a more convenient form that allows for an arbitrary filtration is given). In this section, we will present an extension of this result that gives a variational representation for positive functionals of both \mathbf{B}^n and \mathbf{X}^n .

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$ be a \mathbf{P} -complete filtered probability space on which are given a d -dimensional standard \mathcal{F}_t -Brownian motion B and an \mathcal{F}_0 -measurable random variable X , which takes values in a Polish space \mathbb{S} and has law ρ . Main result of this section is Proposition 4.1 which gives a variational representation for $-\log \mathbf{E}[\exp(-f(B, X))]$, where $f \in \mathbb{M}_b(\mathcal{C} \times \mathbb{S})$. The proof is given in Section 9. We will apply this result in Sections 5 – 7 for the setting where d is replaced by nd , \mathbb{S} by \mathbb{R}_+^n , and ρ by $\theta^{\otimes n}$.

Note that since X is \mathcal{F}_0 -measurable, B and X are independent. Consider the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$, where $\tilde{\Omega} = \mathcal{C}$, $\tilde{\mathcal{F}} = \mathcal{B}(\mathcal{C})$ and $\tilde{\mathbf{P}}$ is the d -dimensional Wiener measure. Under $\tilde{\mathbf{P}}$ the canonical coordinate process $\tilde{W} \doteq \{\tilde{W}_t(\tilde{\omega}) \doteq \tilde{\omega}(t), 0 \leq t \leq T\}$ is a standard d -dimensional Brownian motion with respect to the filtration $\{\tilde{\mathcal{F}}_t^{\tilde{W}}\} \doteq \{\sigma\{\tilde{W}(s) : s \leq t\}\}$. Let $\{\tilde{\mathcal{F}}_t\}$ be the augmented filtration, namely $\tilde{\mathcal{F}}_t \doteq \sigma\{\tilde{\mathcal{F}}_t^{\tilde{W}} \cup \tilde{\mathcal{N}}\}$ and $\tilde{\mathcal{N}}$ is the collection of all $\tilde{\mathbf{P}}$ -null sets. For $f \in \mathbb{M}_b(\mathcal{C} \times \mathbb{S})$, define

$$\tilde{f}(x) \doteq -\log \tilde{\mathbf{E}} \left[\exp \left(-f(\tilde{W}, x) \right) \right], \quad x \in \mathbb{S}. \quad (4.2)$$

It follows from the independence between B and X that

$$-\log \mathbf{E}[\exp(-f(B, X))] = -\log \mathbf{E} \left[\exp \left(-\tilde{f}(X) \right) \right].$$

Applying classical results of Donsker–Varadhan (cf. [10] Proposition 1.4.2) to RHS, we have the following representation formula from the above equality

$$-\log \mathbf{E}[\exp(-f(B, X))] = \inf_{\Pi \in \mathcal{P}(\mathbb{S})} \left[R(\Pi \| \rho) + \int_{\mathbb{S}} \tilde{f}(x) \Pi(dx) \right]. \quad (4.3)$$

Consider the collection of processes

$$\tilde{\mathcal{A}} \doteq \{\tilde{\psi} : \text{the process } \tilde{\psi} \text{ is } \tilde{\mathcal{F}}_t\text{-progressively measurable and } \tilde{\mathbf{E}} \int_0^T \|\tilde{\psi}(s)\|^2 ds < \infty\}.$$

From Theorem 3.1 in [2] we now have the following variational formula

$$\tilde{f}(x) = \inf_{\tilde{\psi} \in \tilde{\mathcal{A}}} \tilde{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 ds + f \left(\tilde{W} + \int_0^\cdot \tilde{\psi}(s) ds, x \right) \right], \quad (4.4)$$

which together with (4.3) gives

$$\begin{aligned} & -\log \mathbf{E}[\exp(-f(B, X))] \\ &= \inf_{\Pi \in \mathcal{P}(\mathbb{S})} \left\{ R(\Pi \| \rho) \right. \\ & \quad \left. + \int_{\mathbb{S}} \inf_{\tilde{\psi} \in \tilde{\mathcal{A}}} \tilde{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\tilde{\psi}(s)\|^2 ds + f \left(\tilde{W} + \int_0^\cdot \tilde{\psi}(s) ds, x \right) \right] \Pi(dx) \right\}. \end{aligned} \quad (4.5)$$

The above representation is inconvenient to use in our large deviation proofs due to the infimum under the integral on the right side. This is taken care of in the following proposition that roughly says that the integral and the infimum can be interchanged and furthermore the

minimal filtration $\tilde{\mathcal{F}}_t$ can be replaced by a general filtration $\{\hat{\mathcal{F}}_t\}$. The fact that one can allow for a general filtration will be crucial in the proof of the lower bound (see Section 7).

Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ be a probability space on which are given a d -dimensional standard Brownian motion \bar{W} and an \mathbb{S} -valued random variable \bar{X} , which is independent of \bar{W} , with law Π . Let $\{\hat{\mathcal{F}}_t\}$ be *any filtration* satisfying the usual conditions, such that \bar{W} is still a standard d -dimensional Brownian motion with respect to $\{\hat{\mathcal{F}}_t\}$ and \bar{X} is $\hat{\mathcal{F}}_0$ -measurable. One example of such a filtration is $\{\bar{\mathcal{F}}_t \doteq \sigma\{\bar{\mathcal{F}}_t^{\bar{W}, \bar{X}} \cup \bar{\mathcal{N}}\}\}$, where $\bar{\mathcal{N}}$ is the collection of all $\bar{\mathbf{P}}$ -null sets and $\bar{\mathcal{F}}_t^{\bar{W}, \bar{X}} \doteq \sigma\{\bar{X}, \bar{W}(s) : s \leq t\}$. Let $\Upsilon_\Pi \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \bar{\mathbf{P}})$ and consider the following collection of processes

$$\begin{aligned} \mathcal{A}(\Upsilon_\Pi) &\doteq \{\hat{\psi} : \text{the process } \hat{\psi} \text{ is } \hat{\mathcal{F}}_t\text{-progressively measurable} \\ &\text{and } \bar{\mathbf{E}} \int_0^T \|\hat{\psi}(s)\|^2 ds < \infty\}. \end{aligned}$$

Proposition 4.1. *Let $f \in \mathbb{M}_b(\mathcal{C} \times \mathbb{S})$. Then*

$$\begin{aligned} &-\log \mathbf{E} [\exp(-f(B, X))] \\ &= \inf_{\Pi, \Upsilon_\Pi} \inf_{\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)} \left\{ R(\Pi \| \rho) + \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 ds + f\left(\bar{W} + \int_0^\cdot \hat{\psi}(s) ds, \bar{X}\right) \right] \right\}, \end{aligned} \quad (4.6)$$

where the outer infimum is over all $\Pi \in \mathcal{P}(\mathbb{S})$ and all systems Υ_Π .

5. Variational representation and tightness properties

In this section we will apply Proposition 4.1 to establish a variational representation for

$$-\frac{1}{n} \log \mathbf{E} [\exp(-nF(\mu^n))],$$

where $F \in \mathbb{C}_b(\mathcal{D})$.

Given $\Pi^n \in \mathcal{P}(\mathbb{R}_+^n)$ such that $R(\Pi^n \| \theta^{\otimes n}) < \infty$, let $\Upsilon_{\Pi^n} \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}_{0 \leq t \leq T}, \bar{\mathbf{P}})$ be a filtered probability space on which one is given an \mathbb{R}_+^n -valued $\hat{\mathcal{F}}_0$ -measurable random variable $\mathbf{S}^n \doteq (\mathbf{S}_i^n)_{i=1}^n$ with law Π^n , and an nd -dimensional standard $\hat{\mathcal{F}}_t$ -Brownian motion $\beta^n \doteq (\beta_i^n)_{i=1}^n$.

Denote by $\bar{\mathbf{E}}$ the expectation under $\bar{\mathbf{P}}$. We can disintegrate Π^n as

$$\Pi^n(dx_1, \dots, dx_n) = \Pi_1^n(dx_1) \Pi_2^n(x_1, dx_2) \cdots \Pi_n^n(x_1, \dots, x_{n-1}, dx_n) \doteq \prod_{i=1}^n \nu_i^n(\mathbf{x}^n, dx_i),$$

where $\mathbf{x}^n \doteq (x_i)_{i=1}^n \in \mathbb{R}_+^n$. Define random measures

$$\nu_i^n \doteq \nu_i^n(\mathbf{S}^n, \cdot), \quad i = 1, \dots, n. \quad (5.1)$$

Let

$$\mathcal{A}_n(\Upsilon_{\Pi^n}) \doteq \{\psi^n \doteq (\psi_i^n)_{i=1}^n : \text{the process } \psi_i^n \text{ is } \mathbb{R}^d\text{-valued}$$

$$\hat{\mathcal{F}}_t\text{-progressively measurable and } \bar{\mathbf{E}} \sum_{i=1}^n \int_0^T \|\psi_i^n(s)\|^2 ds < \infty\}.$$

For $\psi^n \in \mathcal{A}_n(\Upsilon_{\Pi^n})$, consider the following controlled processes

$$\tilde{\beta}^n \doteq (\tilde{\beta}_i^n)_{i=1}^n, \quad \tilde{\beta}_i^n(t) \doteq \beta_i^n(t) + \int_0^t \psi_i^n(s) ds, \quad i = 1, \dots, n, \quad (5.2)$$

$$\tilde{\mu}^n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{\tilde{\beta}_i^n(t)} \mathbf{1}_{\{S_i^n > \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}}. \quad (5.3)$$

Note that since $R(\Pi^n \|\theta^{\otimes n}) < \infty$, Π^n is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}_+^n . Hence a.s. $\bar{\mathbf{P}}$ we can enumerate $\{S_i^n\}_{i=1}^n$ in a strictly increasing order, and the unique solution of (5.3) can be written explicitly in a recursive manner. From Proposition 4.1 we have the following variational formula which will be the starting point of our proof to Theorem 3.2.

Lemma 5.1. *Let $F \in \mathbb{C}_b(\mathcal{D})$. Then*

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))] \\ &= \inf_{\Pi^n, \Upsilon_{\Pi^n}} \inf_{\psi^n \in \mathcal{A}_n(\Upsilon_{\Pi^n})} \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\nu_i^n \|\theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 ds \right) + F(\tilde{\mu}^n) \right], \end{aligned} \quad (5.4)$$

where the outside infimum is over all $\Pi^n \in \mathcal{P}(\mathbb{R}_+^n)$ and all systems Υ_{Π^n} .

Proof. Note that

$$F(\mu^n) = F \left(\left\{ \frac{1}{n} \sum_{i=1}^n \delta_{B_i(t)} \mathbf{1}_{\{X_i > \int_0^t \langle \zeta, \mu^n(s) \rangle ds\}} \right\}_{t \in [0, T]} \right) \doteq \Psi(\mathbf{B}^n, \mathbf{X}^n)$$

for some $\Psi \in \mathbb{M}_b(\mathcal{C}^n \times \mathbb{R}_+^n)$, where $\mathbf{B}^n \doteq (B_i)_{i=1}^n$ and $\mathbf{X}^n \doteq (X_i)_{i=1}^n$. The function Ψ depends also on n but that dependence is suppressed in the notation. With the same measurable function Ψ , $F(\tilde{\mu}^n) = \Psi(\tilde{\beta}^n, \mathbf{S}^n)$ a.s. Applying Proposition 4.1 with $d, \mathbb{S}, f, \bar{W}, \bar{X}, \mathcal{C}$ and ρ replaced by $nd, \mathbb{R}_+^n, n\Psi, \beta^n, \mathbf{S}^n, \mathcal{C}^n$ and $\theta^{\otimes n}$ respectively, we have

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{E} [\exp (-nF(\mu^n))] \\ &= \inf_{\Pi^n, \Upsilon_{\Pi^n}} \inf_{\psi^n \in \mathcal{A}_n(\Upsilon_{\Pi^n})} \left\{ \frac{1}{n} R(\Pi^n \|\theta^{\otimes n}) + \bar{\mathbf{E}} \left[\frac{1}{2n} \sum_{i=1}^n \int_0^T \|\psi_i^n(s)\|^2 ds + F(\tilde{\mu}^n) \right] \right\}. \end{aligned}$$

Using chain rule for relative entropies (cf. [10], Theorem C.3.1) it follows that

$$R(\Pi^n \|\theta^{\otimes n}) = \bar{\mathbf{E}} \left[\sum_{i=1}^n R(\nu_i^n \|\theta) \right],$$

where $(\nu_i^n)_{i=1}^n$ is defined in (5.1). The result follows on combining the above two displays. \square

We now present some tightness results that will play a key role in our proofs. For each $n \in \mathbb{N}$, let Π^n , Υ_{Π^n} (with the associated \mathcal{S}^n) and $\psi^n \in \mathcal{A}_n(\Upsilon_{\Pi^n})$ be such that

$$\sup_{n \in \mathbb{N}} \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\nu_i^n \| \theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 ds \right) \right] \doteq C_0 < \infty. \quad (5.5)$$

For $i = 1, \dots, n$, define \mathcal{R}^W -valued random variables ρ_i^n as

$$\rho_i^n(A \times [0, t]) \doteq \int_0^t \delta_A(\psi_i^n(s)) ds, \quad A \in \mathcal{B}(\mathbb{R}^d), \quad t \in [0, T].$$

Recall the space Ξ from Section 3.1. Define a sequence $\{(Q^n, \nu^n)\}_{n \in \mathbb{N}}$ of $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}_+)$ -valued random variables as

$$Q^n(A) \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(\beta_i^n, \beta_i^n, \rho_i^n, \mathcal{S}_i^n)}(A), \quad A \in \mathcal{B}(\Xi) \quad (5.6)$$

and

$$\nu^n(A) \doteq \frac{1}{n} \sum_{i=1}^n \nu_i^n(A), \quad A \in \mathcal{B}(\mathbb{R}_+). \quad (5.7)$$

We then have the following tightness result. In proving this result we will use some basic facts about tightness functions. Recall (see [10, Section 5.3]) that a measurable map ψ from a Polish space \mathbb{S} to \mathbb{R} is called a tightness function if it is bounded from below and has pre-compact level sets (i.e. for some $c \in \mathbb{R}$, $\psi(x) \geq c$ for all $x \in \mathbb{S}$ and the closure of $\{x \in \mathbb{S} : \psi(x) \leq m\}$ is compact for all $m \in \mathbb{R}$). The following two facts about tightness functions will be used:

- If ψ is a tightness function on \mathbb{S} and $\{\gamma_n\} \subset \mathcal{P}(\mathbb{S})$ is such that $\sup_n \int_{\mathbb{S}} \psi(x) \gamma_n(dx) < \infty$, then $\{\gamma_n\}$ is tight.
- If ψ is a tightness function on \mathbb{S} , then $\Psi : \mathcal{P}(\mathbb{S}) \rightarrow \mathbb{R} \cup \{\infty\}$ defined as $\Psi(\gamma) \doteq \int_{\mathbb{S}} \psi(x) \gamma(dx)$ is a tightness function on $\mathcal{P}(\mathbb{S})$.

We will also appeal to the facts that a sequence $\{\Gamma^n\}$ of $\mathcal{P}(\mathbb{S})$ -valued random variables is tight if and only if $\gamma^n \doteq \mathbf{E}\Gamma^n$ is a tight sequence in $\mathcal{P}(\mathbb{S})$ (see [23, Chapter 1]); and for Polish spaces \mathbb{S}_i , $i = 1, 2$, a sequence of probability measures on $\mathbb{S}_1 \times \mathbb{S}_2$ is tight if the corresponding sequences of marginal distributions on \mathbb{S}_i , $i = 1, 2$, are tight.

Lemma 5.2. *Suppose (5.5) holds. The sequence of random variables $\{(Q^n, \nu^n)\}_{n \geq 1}$ is tight in $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}_+)$.*

Proof. We first argue tightness of $\{Q^n\}$. It suffices to show that the sequences $\{Q_{(i)}^n\}$ are tight for $i = 1, 2, 3, 4$. Since $\{\beta_i^n\}$ are standard Brownian motions, $\{Q_{(2)}^n\}$ is clearly tight. Next we consider $Q_{(3)}^n$. Note that

$$g_1(r) \doteq \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 r(dy dt), \quad r \in \mathcal{R}^W,$$

is a tightness function on \mathcal{R}^W . This says that $G_1 : \mathcal{P}(\mathcal{R}^W) \rightarrow [0, \infty]$ defined as

$$G_1(m) \doteq \int_{\mathcal{R}^W} g_1(r) m(dr), \quad m \in \mathcal{P}(\mathcal{R}^W)$$

is a tightness function on $\mathcal{P}(\mathcal{R}^W)$. Next note that

$$\begin{aligned}\bar{\mathbf{E}}G_1(Q_{(3)}^n) &= \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}}g_1(\rho_i^n) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho_i^n(dy dt) \\ &= \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \int_0^T \|\psi_i^n(s)\|^2 ds \leq 2C_0.\end{aligned}$$

This proves that $\{Q_{(3)}^n\}$ is tight. We now argue tightness of $\{Q_{(1)}^n\}$. Define a sequence $\{\tilde{Q}^n\}_{n \in \mathbb{N}}$ of $\mathcal{P}(\mathcal{C})$ -valued random variables as

$$\tilde{Q}^n(A) \doteq \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{u}_i^n}(A), \quad A \in \mathcal{B}(\mathcal{C}),$$

where

$$\mathbf{u}_i^n(t) \doteq \int_{\mathbb{R}^d \times [0, t]} y \rho_i^n(dy ds) = \int_0^t \psi_i^n(s) ds, \quad t \in [0, T].$$

We claim that $\{\tilde{Q}^n\}_{n \in \mathbb{N}}$ is a tight sequence. To see this, note that $g_2: \mathcal{C} \rightarrow [0, \infty]$ defined as

$$g_2(x) \doteq \begin{cases} \int_0^T \|\dot{x}(s)\|^2 ds + \|x(0)\|^2, & \text{if } x \text{ is absolutely continuous,} \\ \infty, & \text{otherwise,} \end{cases}$$

is a tightness function on \mathcal{C} , from which it follows that $G_2: \mathcal{P}(\mathcal{C}) \rightarrow [0, \infty]$, defined as

$$G_2(m) \doteq \int_{\mathcal{C}} g_2(x) m(dx), \quad m \in \mathcal{P}(\mathcal{C}),$$

is a tightness function on $\mathcal{P}(\mathcal{C})$. Also,

$$\bar{\mathbf{E}}(G_2(\tilde{Q}^n)) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}}g_2(\mathbf{u}_i^n) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \int_0^T \|\psi_i^n(s)\|^2 ds \leq 2C_0.$$

This proves tightness of $\{\tilde{Q}^n\}_{n \in \mathbb{N}}$. Define the sequence $\{\bar{Q}_{(2,3)}^n\}_{n \in \mathbb{N}}$ of $\mathcal{P}(\mathcal{C} \times \mathcal{C})$ -valued random variables as

$$\bar{Q}_{(2,3)}^n(A) \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(\beta_i^n, \mathbf{u}_i^n)}(A), \quad A \in \mathcal{B}(\mathcal{C} \times \mathcal{C}).$$

Then, from tightness of $\{Q_{(2)}^n\}$ and $\{\tilde{Q}^n\}$ it follows that $\{\bar{Q}_{(2,3)}^n\}$ is tight. Next, noting that the map $g_3: \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ defined as $g_3(x, u) \doteq x + u$ is continuous and that $Q_{(1)}^n = \bar{Q}_{(2,3)}^n \circ g_3^{-1}$, we get tightness of $\{Q_{(1)}^n\}$.

Finally, tightness of $\{Q_{(4)}^n\}$ and $\{\nu^n\}$ can be proved using a standard argument as in the proof of Sanov's theorem (see e.g. [10, Theorem 2.2.1]). However we provide details for sake of completeness. Note that $G_3(\cdot) \doteq R(\cdot \|\theta)$ is a convex tightness function on $\mathcal{P}(\mathbb{R}_+)$ (cf. [10, Lemma 1.4.3]), and that by (5.5) and Jensen's inequality,

$$\bar{\mathbf{E}}G_3(\nu^n) = \bar{\mathbf{E}}R(\nu^n \|\theta) \leq \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n R(\nu_i^n \|\theta) \right] \leq C_0.$$

So $\{\nu^n\}$ is tight, which implies that $\{\bar{\mathbf{E}}\nu^n\}$ is also tight. Next, since ν_i^n is the conditional distribution used to select \mathbf{S}_i^n , for any $f \in \mathbb{M}_b(\mathbb{R}^d)$,

$$\bar{\mathbf{E}} \int_{\mathbb{R}^d} f(x) Q_{(4)}^n(dx) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} f(\mathbf{S}_i^n) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \int_{\mathbb{R}^d} f(x) \nu_i^n(dx) = \bar{\mathbf{E}} \int_{\mathbb{R}^d} f(x) \nu^n(dx).$$

So $\bar{\mathbf{E}}Q_{(4)}^n = \bar{\mathbf{E}}\nu^n$, from which we have the desired tightness of $\{Q_{(4)}^n\}$. \square

The following lemma gives a useful characterization of weak limit points of $\{(Q^n, \nu^n) : n \in \mathbb{N}\}$. Recall the space \mathcal{P}_∞ introduced in Section 3.2.

Lemma 5.3. *Suppose (5.5) holds and suppose (Q^n, ν^n) converges along a subsequence, in distribution, to (Q^*, ν^*) given on some probability space $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$. Then $Q_{(4)}^* = \nu^*$, $Q^* \in \mathcal{P}_\infty$, and $\mathcal{J}(Q^*) < \infty$, a.s. \mathbf{P}^* .*

Proof. The first statement, namely $Q_{(4)}^* = \nu^*$ a.s. \mathbf{P}^* , can be proved using a standard martingale argument as in the proof of Sanov's theorem (see e.g. [10, Theorem 2.2.1]). We omit the details.

In order to check the second statement that $Q^* \in \mathcal{P}_\infty$ a.s. \mathbf{P}^* , we need to verify that, a.s., $Q_{(4)}^* \ll \theta$ and under Q^* , properties (1)–(3) of Section 3.2 are satisfied (with Θ replaced by Q^*). Without loss of generality assume that (Q^n, ν^n) converges weakly to (Q^*, ν^*) along the whole sequence. Recall the canonical variables $(\tilde{\mathbf{b}}, \mathbf{b}, \boldsymbol{\rho}, \tilde{\boldsymbol{\sigma}})$ and the canonical filtration $\{\mathcal{G}_t\}$ introduced in Section 3.1. We can find a countable collection $\{\eta_j\}_{j=1}^\infty$ of continuous nonnegative functions with compact support in \mathbb{R}^d , such that, denoting

$$\boldsymbol{\rho}_t^{\eta_j}(\xi) \doteq \int_{\mathbb{R}^d \times [0, t]} \eta_j(y) \rho(dy ds), \quad t \in [0, T], \quad j \in \mathbb{N}, \quad \xi = (\tilde{\mathbf{b}}, \mathbf{b}, \boldsymbol{\rho}, \tilde{\boldsymbol{\sigma}}) \in \Xi$$

and defining the stochastic process

$$\mathcal{E}(t) \doteq (\tilde{\mathbf{b}}_t, \mathbf{b}_t, (\boldsymbol{\rho}_t^{\eta_j})_{j \in \mathbb{N}}, \tilde{\boldsymbol{\sigma}}), \quad t \in [0, T]$$

with sample paths in $\mathcal{C}_\infty \doteq \mathbb{C}([0, T] : \mathbb{R}^\infty)$, we have $\mathcal{G}_t = \sigma\{\mathcal{E}(\cdot \wedge t)\}$ for $t \in [0, T]$.

We first verify (1). It suffices to check that for all $0 \leq s \leq t \leq T$, $g \in \mathbb{C}_b^2(\mathbb{R}^d)$ and $f \in \mathbb{C}_b(\mathcal{C}_\infty)$,

$$\mathbf{E}^* \left| \mathbf{E}^{Q^*} f(\mathcal{E}(\cdot \wedge s)) \left(g(\mathbf{b}_t) - g(\mathbf{b}_s) - \frac{1}{2} \int_s^t \Delta g(\mathbf{b}_u) du \right) \right|^2 = 0. \quad (5.8)$$

Since Q^n converges weakly to Q^* the left side equals

$$\begin{aligned} & \lim_{n \rightarrow \infty} \bar{\mathbf{E}} \left| \mathbf{E}^{Q^n} f(\mathcal{E}(\cdot \wedge s)) \left(g(\mathbf{b}_t) - g(\mathbf{b}_s) - \frac{1}{2} \int_s^t \Delta g(\mathbf{b}_u) du \right) \right|^2 \\ &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \bar{\mathbf{E}} \left| \sum_{i=1}^n f(\mathcal{E}_i^n(\cdot \wedge s)) \left(g(\boldsymbol{\beta}_i^n(t)) - g(\boldsymbol{\beta}_i^n(s)) - \frac{1}{2} \int_s^t \Delta g(\boldsymbol{\beta}_i^n(u)) du \right) \right|^2, \end{aligned}$$

where \mathcal{E}_i^n is defined similarly to \mathcal{E} by replacing $(\tilde{\mathbf{b}}, \mathbf{b}, \boldsymbol{\rho}, \tilde{\boldsymbol{\sigma}})$ with $(\tilde{\boldsymbol{\beta}}_i^n, \boldsymbol{\beta}_i^n, \boldsymbol{\rho}_i^n, \mathbf{S}_i^n)$. Conditioning on $\hat{\mathcal{F}}_s$ and using the fact that $\boldsymbol{\beta}^n$ is a standard $\hat{\mathcal{F}}_t$ -Brownian motion, we see that cross product terms do not contribute when the above squared sum is written as a double sum. So the above limit is 0 which proves (5.8).

Consider now (2). It suffices to show that

$$\mathbf{E}^* \mathbf{E}^{Q^*} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy dt) < \infty. \quad (5.9)$$

Note that $Q^n \Rightarrow Q^*$ as $n \rightarrow \infty$ and

$$\bar{\mathbf{E}} \mathbf{E}^{Q^n} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy dt) = \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \int_0^T \|\psi_i^n(t)\|^2 dt \leq 2C_0.$$

Thus in order to prove (5.9), it suffices to show that the function

$$Q \mapsto \mathbf{E}^Q \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy dt) \quad (5.10)$$

is a lower semi-continuous function from $\mathcal{P}(\Xi)$ to $[0, \infty]$. This is immediate from Fatou's lemma on observing that the function

$$\xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \mapsto \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds)$$

is a lower semi-continuous function from Ξ to $[0, \infty]$. This proves (2).

Next we verify (3). It suffices to show that for each $t \in [0, T]$

$$\mathbf{E}^* \mathbf{E}^{Q^*} \left(\left| \tilde{\mathbf{b}}_t - \mathbf{b}_t - \int_{\mathbb{R}^d \times [0, t]} y \boldsymbol{\rho}(dy ds) \right| \wedge 1 \right) = 0. \quad (5.11)$$

Note that for each $t \in [0, T]$, the function

$$\xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \mapsto \left| \tilde{b}_t - b_t - \int_{\mathbb{R}^d \times [0, t]} y \rho(dy ds) \right| \wedge 1 \quad (5.12)$$

is a continuous and bounded function from Ξ to \mathbb{R}_+ . Also, from (5.2),

$$\begin{aligned} & \bar{\mathbf{E}} \mathbf{E}^{Q^n} \left(\left| \tilde{\mathbf{b}}_t - \mathbf{b}_t - \int_{\mathbb{R}^d \times [0, t]} y \boldsymbol{\rho}(dy ds) \right| \wedge 1 \right) \\ &= \frac{1}{n} \sum_{i=1}^n \bar{\mathbf{E}} \left(\left| \tilde{\boldsymbol{\beta}}_i^n(t) - \boldsymbol{\beta}_i^n(t) - \int_0^t \boldsymbol{\psi}_i^n(s) ds \right| \wedge 1 \right) = 0. \end{aligned}$$

Combining the above two observations and recalling that $Q^n \Rightarrow Q^*$ we have (5.11).

Next we show that $Q_{(4)}^* \ll \theta$ a.s. By lower semi-continuity and convexity of the function $R(\cdot \|\theta)$, we have on using the first statement of the lemma that

$$\mathbf{E}^* R(Q_{(4)}^* \|\theta) = \mathbf{E}^* R(\nu^* \|\theta) \leq \liminf_{n \rightarrow \infty} \bar{\mathbf{E}} R(\nu^n \|\theta) \leq \liminf_{n \rightarrow \infty} \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n R(\nu_i^n \|\theta) \right] \leq C_0. \quad (5.13)$$

So $R(Q_{(4)}^* \|\theta) < \infty$ and consequently $Q_{(4)}^* \ll \theta$ a.s. \mathbf{P}^* .

Finally, the third statement, namely $\mathcal{J}(Q^*) < \infty$ a.s., follows immediately on combining (5.9) and (5.13). \square

For the following two lemmas, define

$$\|\tilde{\beta}^n\|_{*,n}^2 \doteq \sup_{0 \leq t \leq T} \sum_{i=1}^n \|\tilde{\beta}_i^n(t)\|^2.$$

From Cauchy–Schwarz and Doob’s inequalities we have for all $n \in \mathbb{N}$

$$\begin{aligned} \frac{1}{n} \bar{\mathbf{E}} \|\tilde{\beta}^n\|_{*,n}^2 &\leq \frac{2}{n} \bar{\mathbf{E}} \sup_{0 \leq t \leq T} \sum_{i=1}^n \left(\|\beta_i^n(t)\|^2 + \left\| \int_0^t \psi_i^n(s) ds \right\|^2 \right) \\ &\leq 8Td + \frac{2T}{n} \bar{\mathbf{E}} \sum_{i=1}^n \int_0^T \|\psi_i^n(s)\|^2 ds \\ &\leq 8Td + 4TC_0 \doteq \tilde{C}_0, \end{aligned} \quad (5.14)$$

where the last line follows from (5.5). Using this moment bound we will now argue tightness of $\{\tilde{\mu}^n\}$.

Lemma 5.4. *Suppose (5.5) holds. The sequence of random variables $\{\tilde{\mu}^n\}_{n \geq 1}$ is tight in \mathcal{D} .*

Proof. We first argue tightness of $\{\tilde{\mu}^n(t)\}$ in $\mathcal{M}(\mathbb{R}^d)$ for each $t \in [0, T]$. Note that

$$G_4(m) \doteq \int_{\mathbb{R}^d} \|x\|^2 m(dx), \quad m \in \mathcal{M}(\mathbb{R}^d)$$

is a tightness function on $\mathcal{M}(\mathbb{R}^d)$. Also from (5.14) we have

$$\bar{\mathbf{E}} G_4(\tilde{\mu}^n(t)) = \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \|\tilde{\beta}_i^n(t)\|^2 \mathbf{1}_{\{S_i^n > \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} \right] \leq \frac{1}{n} \bar{\mathbf{E}} \|\tilde{\beta}^n\|_{*,n}^2 \leq \tilde{C}_0.$$

Thus $\{\tilde{\mu}^n(t)\}$ is tight in $\mathcal{M}(\mathbb{R}^d)$ for each $t \in [0, T]$.

Next we consider fluctuations of $\tilde{\mu}^n$. Recall that $\tilde{\mu}^n$ is defined on some system Υ_{Π^n} . For $\delta \in [0, T]$, let $\mathcal{T}^{\delta,n}$ be the collection of all $[0, T - \delta]$ -valued stopping times τ on Υ_{Π^n} . In order to argue tightness of $\{\tilde{\mu}^n\}$ in \mathcal{D} , by Aldous–Kurtz tightness criterion (cf. [19, Theorem 2.7]), it suffices to show that

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}^{\delta,n}} \bar{\mathbf{E}} d_{BL}(\tilde{\mu}^n(\tau + \delta), \tilde{\mu}^n(\tau)) = 0. \quad (5.15)$$

Note that

$$\begin{aligned}
& d_{BL}(\tilde{\mu}^n(\tau + \delta), \tilde{\mu}^n(\tau)) \\
&= \sup_{\|f\|_{BL} \leq 1} |\langle f, \tilde{\mu}^n(\tau + \delta) \rangle - \langle f, \tilde{\mu}^n(\tau) \rangle| \\
&\leq \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^n \left| f(\tilde{\beta}_i^n(\tau + \delta)) \mathbf{1}_{\{S_i^n > \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} - f(\tilde{\beta}_i^n(\tau)) \mathbf{1}_{\{S_i^n > \int_0^\tau \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} \right| \\
&\leq \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^n \left| f(\tilde{\beta}_i^n(\tau + \delta)) - f(\tilde{\beta}_i^n(\tau)) \right| \mathbf{1}_{\{S_i^n > \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} \\
&\quad + \sup_{\|f\|_{BL} \leq 1} \frac{1}{n} \sum_{i=1}^n \left| f(\tilde{\beta}_i^n(\tau)) \left(\mathbf{1}_{\{S_i^n > \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} - \mathbf{1}_{\{S_i^n > \int_0^\tau \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} \right) \right| \\
&\doteq \mathcal{T}_1^n + \mathcal{T}_2^n.
\end{aligned}$$

It then suffices to show that for $j = 1, 2$,

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}^{\delta, n}} \bar{\mathbf{E}} \mathcal{T}_j^n = 0. \quad (5.16)$$

For \mathcal{T}_1^n , we have

$$\begin{aligned}
\mathcal{T}_1^n &\leq \frac{1}{n} \sum_{i=1}^n \left\| \tilde{\beta}_i^n(\tau + \delta) - \tilde{\beta}_i^n(\tau) \right\| \\
&\leq \frac{1}{n} \sum_{i=1}^n \left(\left\| \beta_i^n(\tau + \delta) - \beta_i^n(\tau) \right\| + \int_\tau^{\tau+\delta} \left\| \psi_i^n(s) \right\| ds \right).
\end{aligned}$$

So by Cauchy–Schwarz inequality,

$$\begin{aligned}
& \sup_{\tau \in \mathcal{T}^{\delta, n}} \bar{\mathbf{E}} \mathcal{T}_1^n \\
&\leq \sup_{\tau \in \mathcal{T}^{\delta, n}} \left(\bar{\mathbf{E}} \frac{1}{n} \sum_{i=1}^n \left\| \beta_i^n(\tau + \delta) - \beta_i^n(\tau) \right\|^2 \right)^{\frac{1}{2}} + \sqrt{\delta} \left(\mathbf{E} \frac{1}{n} \sum_{i=1}^n \int_0^T \left\| \psi_i^n(s) \right\|^2 ds \right)^{\frac{1}{2}} \\
&\leq \sqrt{d\delta} + \sqrt{2C_0\delta},
\end{aligned}$$

which implies that (5.16) holds for $j = 1$.

Now consider \mathcal{T}_2^n . Fix $K \in (0, \infty)$. Since $\|f\|_\infty \leq 1$,

$$\begin{aligned}
\mathcal{T}_2^n &\leq \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\left\{ \int_0^\tau \langle \zeta, \tilde{\mu}^n(s) \rangle ds < S_i^n \leq \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds \right\}} \\
&\leq \mathbf{1}_{\{\|\tilde{\beta}^n\|_{*,n}^2 > nK\}} + \mathbf{1}_{\{\|\tilde{\beta}^n\|_{*,n}^2 \leq nK\}} Q_{(4)}^n \left(\left(\int_0^\tau \langle \zeta, \tilde{\mu}^n(s) \rangle ds, \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds \right) \right) \\
&\doteq \mathcal{T}_{2,1}^n + \mathcal{T}_{2,2}^n.
\end{aligned}$$

For $\mathcal{T}_{2,1}^n$, from (5.14) it follows that

$$\bar{\mathbf{E}}\mathcal{T}_{2,1}^n \leq \frac{\bar{\mathbf{E}}\|\tilde{\beta}^n\|_{*,n}^2}{nK} \leq \frac{\tilde{C}_0}{K}. \quad (5.17)$$

For $\mathcal{T}_{2,2}^n$, note that on the set $\{\|\tilde{\beta}^n\|_{*,n}^2 \leq nK\}$, we have from (2.2) that for $t \in [0, T]$,

$$\begin{aligned} \langle \zeta, \tilde{\mu}^n(t) \rangle &\leq \frac{1}{n} \sum_{i=1}^n \zeta(\tilde{\beta}_i^n(t)) \leq \frac{1}{n} \sum_{i=1}^n \tilde{C}_\zeta (1 + \|\tilde{\beta}_i^n(t)\|^2) \leq \tilde{C}_\zeta + \frac{\tilde{C}_\zeta \|\tilde{\beta}^n\|_{*,n}^2}{n} \\ &\leq \tilde{C}_\zeta (1 + K), \end{aligned} \quad (5.18)$$

and hence

$$\left| \int_0^{\tau+\delta} \langle \zeta, \tilde{\mu}^n(s) \rangle ds - \int_0^\tau \langle \zeta, \tilde{\mu}^n(s) \rangle ds \right| \leq C_K \delta$$

with $C_K = \tilde{C}_\zeta(1 + K)$. From this it follows that

$$\mathcal{T}_{2,2}^n \leq \mathbf{1}_{\{\|\tilde{\beta}^n\|_{*,n}^2 \leq nK\}} \max_{j \in \mathbb{N}_0} Q_{(4)}^n([jC_K\delta, (j+2)C_K\delta]).$$

Using this we have

$$\sup_{\tau \in \mathcal{T}^{\delta,n}} \bar{\mathbf{E}}\mathcal{T}_{2,2}^n \leq \bar{\mathbf{E}} \max_{j \in \mathbb{N}_0} Q_{(4)}^n([j(C_K\delta, (j+2)C_K\delta])). \quad (5.19)$$

By Lemmas 5.2 and 5.3, we can assume without loss of generality that (Q^n, ν^n) converges weakly along the whole sequence to (Q^*, ν^*) , given on some probability space $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$, with $Q^* \in \mathcal{P}_\infty$ a.s. \mathbf{P}^* . Note that if $\{m_n\}, m$ are probability measures on $[0, \infty)$ such that m is absolutely continuous with respect to the Lebesgue measure, $m_n \rightarrow m$ as $n \rightarrow \infty$, and $\delta, c > 0$, then

$$\max_{j \in \mathbb{N}_0} m_n([j\delta, j\delta + c]) \rightarrow \max_{j \in \mathbb{N}_0} m([j\delta, j\delta + c]). \quad (5.20)$$

It then follows from (5.19), dominated convergence theorem, (5.20) and Remark 3.1 that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}^{\delta,n}} \bar{\mathbf{E}}\mathcal{T}_2^n &\leq \limsup_{n \rightarrow \infty} \bar{\mathbf{E}} \max_{j \in \mathbb{N}_0} Q_{(4)}^n([jC_K\delta, (j+2)C_K\delta]) \\ &= \mathbf{E}^* \max_{j \in \mathbb{N}_0} Q_{(4)}^*([jC_K\delta, (j+2)C_K\delta]). \end{aligned}$$

Using (5.13), for every $M > 1$,

$$\mathbf{E}^* \max_{j \in \mathbb{N}_0} Q_{(4)}^*([jC_K\delta, (j+2)C_K\delta]) \leq 2MC_K\delta + \frac{C_0}{\log M}. \quad (5.21)$$

This shows that the left side in the above equation converges to 0 as $\delta \rightarrow 0$ and $M \rightarrow \infty$. Combining this convergence with (5.17) and then sending $K \rightarrow \infty$ implies that (5.16) also holds for $j = 2$ which completes the proof of tightness of $\{\tilde{\mu}^n\}$. \square

The following lemma gives an important characterization of weak limit points of $\{\tilde{\mu}^n\}$. Recall the definition of ϖ_Θ for $\Theta \in \mathcal{P}_\infty^1$ given in Section 3.2.

Lemma 5.5. *Suppose (5.5) holds and $(Q^n, \nu^n, \tilde{\mu}^n)$ converges along a subsequence, in distribution, to $(Q^*, \nu^*, \tilde{\mu}^*)$ given on some probability space $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$. Then \mathbf{P}^* -a.s., $Q^* \in \mathcal{P}_\infty^1$ and $\tilde{\mu}^* = \varpi_{Q^*}$, i.e.*

$$\langle f, \tilde{\mu}^*(t) \rangle = \mathbf{E}^{Q^*} \left[f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds\}} \right], \quad \forall f \in \mathbb{C}_b(\mathbb{R}^d), t \in [0, T]. \quad (5.22)$$

Proof. Since $R(\Pi^n \|\theta^{\otimes n}) < \infty$, we have that a.s., $\{\mathbf{S}_i^n\}_{i=1}^n$ are distinct and hence, a.s., the total mass of $\tilde{\mu}^n(t)$ decreases by at most $1/n$ at any time instant t . Consequently, $0 \leq \tilde{\mu}^n(t-)(A) - \tilde{\mu}^n(t)(A) \leq 1/n$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ and $t \in [0, T]$, a.s. Also, without loss of generality assume that $(Q^n, \nu^n, \tilde{\mu}^n)$ converges weakly to $(Q^*, \nu^*, \tilde{\mu}^*)$ along the whole sequence. We claim that, jointly with the above collection,

$$\int_0^\cdot \langle \zeta, \tilde{\mu}^n(s) \rangle ds \Rightarrow \int_0^\cdot \langle \zeta, \tilde{\mu}^*(s) \rangle ds \text{ in } \mathbb{C}([0, T] : \mathbb{R}), \quad (5.23)$$

as $n \rightarrow \infty$. Suppose for now that the claim holds. By appealing to Skorokhod representation we can assume without loss of generality that $(Q^n, \nu^n, \tilde{\mu}^n, \int_0^\cdot \langle \zeta, \tilde{\mu}^n(s) \rangle ds)$ converges a.s. on $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$. From this and Lemma 5.3 we can find $\mathcal{N} \in \mathcal{F}^*$ such that $\mathbf{P}^*(\mathcal{N}) = 0$ and on \mathcal{N}^c , we have that $(Q^n, \nu^n, \tilde{\mu}^n, \int_0^\cdot \langle \zeta, \tilde{\mu}^n(s) \rangle ds) \rightarrow (Q^*, \nu^*, \tilde{\mu}^*, \int_0^\cdot \langle \zeta, \tilde{\mu}^*(s) \rangle ds)$, $Q_{(4)}^* = \nu^*$, $Q^* \in \mathcal{P}_\infty$, $\mathcal{J}(Q^*) < \infty$ and $0 \leq \tilde{\mu}^n(t-)(A) - \tilde{\mu}^n(t)(A) \leq 1/n$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ and $t \in [0, T]$. The last statement implies that on \mathcal{N}^c

$$\sup_{t \in (0, 1]} d_{BL}(\tilde{\mu}^n(t), \tilde{\mu}^n(t-)) = \sup_{t \in (0, 1]} \sup_{\|f\|_{BL} \leq 1} |\langle f, \tilde{\mu}^n(t) \rangle - \langle f, \tilde{\mu}^n(t-) \rangle| \leq \frac{1}{n} \rightarrow 0$$

as $n \rightarrow \infty$ and hence $\tilde{\mu}^* \in \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$ on \mathcal{N}^c . Fix $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t \in [0, T]$. Then, on \mathcal{N}^c , $\langle f, \tilde{\mu}^n(t) \rangle \rightarrow \langle f, \tilde{\mu}^*(t) \rangle$. Define for $t \in [0, T]$, $g_t : \mathcal{P}(\Xi) \times \mathcal{D} \rightarrow \mathbb{R}$ as

$$g_t(\Theta, \pi) = \mathbf{E}^\Theta \left[f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi(s) \rangle ds\}} \right]. \quad (5.24)$$

From (5.2) and (5.3) we have $\langle f, \tilde{\mu}^n(t) \rangle = g_t(Q^n, \tilde{\mu}^n)$.

Noting that the RHS of (5.22) is $g_t(Q^*, \tilde{\mu}^*)$, it then suffices to show that as $n \rightarrow \infty$, $g_t(Q^n, \tilde{\mu}^n) \rightarrow g_t(Q^*, \tilde{\mu}^*)$ in probability. Note that

$$|g_t(Q^n, \tilde{\mu}^n) - g_t(Q^*, \tilde{\mu}^*)| \leq |g_t(Q^n, \tilde{\mu}^n) - g_t(Q^n, \tilde{\mu}^*)| + |g_t(Q^n, \tilde{\mu}^*) - g_t(Q^*, \tilde{\mu}^*)|. \quad (5.25)$$

Since on \mathcal{N}^c , $Q^* \in \mathcal{P}_\infty$, by Remark 3.1 we have on this set,

$$\lim_{n \rightarrow \infty} |g_t(Q^n, \tilde{\mu}^*) - g_t(Q^*, \tilde{\mu}^*)| = 0. \quad (5.26)$$

For the first term on the right hand side of (5.25), we have, on \mathcal{N}^c ,

$$\begin{aligned} & |g_t(Q^n, \tilde{\mu}^n) - g_t(Q^n, \tilde{\mu}^*)| \\ & \leq \mathbf{E}^{Q^n} \left| f(\tilde{\mathbf{b}}_t) \left(\mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds\}} - \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds\}} \right) \right| \\ & \leq \|f\|_\infty Q^n \left(\tilde{\sigma} \text{ lies between } \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds \text{ and } \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds \right). \end{aligned}$$

From convergence of $\int_0^\cdot \langle \zeta, \tilde{\mu}^n(s) \rangle ds$ to $\int_0^\cdot \langle \zeta, \tilde{\mu}^*(s) \rangle ds$ and (5.18), it follows that for any $\delta > 0$,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} |g_t(Q^n, \tilde{\mu}^n) - g_t(Q^n, \tilde{\mu}^*)| \\ & \leq \|f\|_\infty \limsup_{n \rightarrow \infty} Q^n \left(\tilde{\sigma} \text{ lies between } \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds \text{ and } \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds \right) \\ & \leq \|f\|_\infty \limsup_{n \rightarrow \infty} \max_{j \in \mathbb{N}_0} Q_{(4)}^n([j\delta, (j+1)\delta]) \\ & = \|f\|_\infty \max_{j \in \mathbb{N}_0} Q_{(4)}^*([j\delta, (j+1)\delta]), \end{aligned}$$

where the last equality uses (5.20). Using a similar bound as in (5.21) and letting $\delta \rightarrow 0$ in the above display gives $\limsup_{n \rightarrow \infty} |g_t(Q^n, \tilde{\mu}^n) - g_t(Q^n, \tilde{\mu}^*)| = 0$ on \mathcal{N}^c . This together with (5.25) and (5.26) proves that on \mathcal{N}^c , for all $t \in [0, T]$, $g_t(Q^n, \tilde{\mu}^n) \rightarrow g_t(Q^*, \tilde{\mu}^*)$ as $n \rightarrow \infty$ and hence $\langle f, \tilde{\mu}^*(t) \rangle = g_t(Q^*, \tilde{\mu}^*)$. Therefore, $\tilde{\mu}^* = \varpi_{Q^*}$ a.s.

Now we prove the claim (5.23). Fix $K \in (0, \infty)$ and define function $\zeta_K(x) \doteq \min\{\zeta(x), K\}$ for $x \in \mathbb{R}^d$. Clearly as $n \rightarrow \infty$,

$$\int_0^\cdot \langle \zeta_K, \tilde{\mu}^n(s) \rangle ds \Rightarrow \int_0^\cdot \langle \zeta_K, \tilde{\mu}^*(s) \rangle ds \text{ in } \mathbb{C}([0, T] : \mathbb{R}). \quad (5.27)$$

From (1.2) it follows that for all $x \in \mathbb{R}^d$,

$$\begin{aligned} \zeta(x) - \zeta_K(x) & \leq \zeta(x) \mathbf{1}_{\{\zeta(x) > K\}} \leq \zeta(x) \left(\frac{\zeta(x)}{K} \right)^{2/p-1} \\ & \leq \frac{C_\zeta^{2/p} (1 + \|x\|^p)^{2/p}}{K^{2/p-1}} \leq \frac{\kappa_1 (1 + \|x\|^2)}{K^{2/p-1}}. \end{aligned} \quad (5.28)$$

So using (5.3) we have

$$\begin{aligned} & \sup_{n \in \mathbb{N}} \sup_{0 \leq t \leq T} \bar{\mathbf{E}} \left| \int_0^t \langle \zeta, \tilde{\mu}^n(s) \rangle ds - \int_0^t \langle \zeta_K, \tilde{\mu}^n(s) \rangle ds \right| \\ & \leq \sup_{n \in \mathbb{N}} \int_0^T \bar{\mathbf{E}} \langle \zeta - \zeta_K, \tilde{\mu}^n(s) \rangle ds \\ & \leq \sup_{n \in \mathbb{N}} \int_0^T \left(\bar{\mathbf{E}} \frac{1}{n} \sum_{i=1}^n \frac{\kappa_1 (1 + \|\tilde{\beta}_i^n(s)\|^2)}{K^{2/p-1}} \right) ds \\ & \leq \sup_{n \in \mathbb{N}} \frac{\kappa_1}{K^{2/p-1}} \left(T + \frac{T}{n} \bar{\mathbf{E}} \|\tilde{\beta}\|_{*,n}^2 \right) \\ & \leq \frac{\kappa_2}{K^{2/p-1}} \\ & \rightarrow 0 \end{aligned}$$

as $K \rightarrow \infty$, where the last inequality follows from (5.14). Also by Fatou's lemma,

$$\begin{aligned} \mathbf{E}^* \left| \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds - \int_0^t \langle \zeta_K, \tilde{\mu}^*(s) \rangle ds \right| & \leq \liminf_{n \rightarrow \infty} \int_0^t \bar{\mathbf{E}} \langle \zeta - \zeta_K, \tilde{\mu}^n(s) \rangle ds \\ & \leq \frac{\kappa_2}{K^{2/p-1}} \\ & \rightarrow 0 \end{aligned}$$

as $K \rightarrow \infty$. The claim (5.23) then follows on combining above two displays with (5.27). \square

6. Laplace upper bound

In this section we prove the *Laplace upper bound* (3.6): For every $F \in \mathbb{C}_b(\mathcal{D})$,

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \bar{\mathbf{E}}[\exp(-nF(\mu^n))] \geq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\}.$$

Let $\varepsilon \in (0, 1)$ be arbitrary. For each $n \in \mathbb{N}$, let $\Pi^n \in \mathcal{P}(\mathbb{R}_+^n)$, Υ_{Π^n} (with the associated \mathbf{S}^n) and $\psi^n \in \mathcal{A}_n(\Upsilon_{\Pi^n})$ be an ε -optimal control in (5.4), namely

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{E}[\exp(-nF(\mu^n))] \\ & \geq \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\nu_i^n \|\theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 ds \right) + F(\tilde{\mu}^n) \right] - \varepsilon, \end{aligned} \quad (6.1)$$

where $(\nu_i^n)_{i=1}^n$ and $\tilde{\mu}^n$ are defined in (5.1) and (5.3). Then with Q^n, ν^n defined as in (5.6) and (5.7), we have from Jensen's inequality that

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{E}[\exp(-nF(\mu^n))] \\ & \geq \bar{\mathbf{E}} \left[R(\nu^n \|\theta) + \mathbf{E}^{Q^n} \left(\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right) + F(\tilde{\mu}^n) \right] - \varepsilon. \end{aligned}$$

The inequality (6.1) implies that

$$\sup_{n \in \mathbb{N}} \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\nu_i^n \|\theta) + \frac{1}{2} \int_0^T \|\psi_i^n(s)\|^2 ds \right) \right] \leq 2\|F\|_\infty + 1,$$

i.e. (5.5) holds with $C_0 = 2\|F\|_\infty + 1$.

Thus from Lemmas 5.2 and 5.4 we have that $\{(Q^n, \nu^n, \tilde{\mu}^n)\}$ is tight. Assume without loss of generality that $(Q^n, \nu^n, \tilde{\mu}^n)$ converges along the whole sequence weakly to $(Q^*, \nu^*, \tilde{\mu}^*)$, given on some probability space $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$. By Lemmas 5.3 and 5.5, we have $Q^* \in \mathcal{P}_\infty^1$, $Q_{(4)}^* = \nu^*$, and $\varpi_{Q^*} = \tilde{\mu}^*$ a.s. \mathbf{P}^* . Using the lower semi-continuity of the function $R(\cdot \|\theta)$ and the map in (5.10), we now have from Fatou's lemma that

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \bar{\mathbf{E}} \left(R(\nu^n \|\theta) + \mathbf{E}^{Q^n} \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right] \right) \\ & \geq \mathbf{E}^* \left(R(\nu^* \|\theta) + \mathbf{E}^{Q^*} \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right] \right) \\ & = \mathbf{E}^*(\mathcal{J}(Q^*)) \\ & \geq \mathbf{E}^*(I(\varpi_{Q^*})). \end{aligned}$$

Also since $F \in \mathbb{C}_b(\mathcal{D})$,

$$\lim_{n \rightarrow \infty} \bar{\mathbf{E}} F(\tilde{\mu}^n) = \mathbf{E}^* F(\tilde{\mu}^*) = \mathbf{E}^* F(\varpi_{Q^*}).$$

Combining above three displays we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \bar{\mathbf{E}} [\exp(-nF(\mu^n))] &\geq \mathbf{E}^*(F(\varpi_{Q^*}) + I(\varpi_{Q^*})) - \varepsilon \\ &\geq \inf_{\pi \in \mathcal{D}} (F(\pi) + I(\pi)) - \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the Laplace upper bound. \square

7. Laplace lower bound

In this section we prove the *Laplace lower bound* (3.7): For every $F \in \mathbb{C}_b(\mathcal{D})$,

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \bar{\mathbf{E}} [\exp(-nF(\mu^n))] \leq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\}.$$

Fix $F \in \mathbb{C}_b(\mathcal{D})$ and let $\varepsilon > 0$ be arbitrary and let $\pi^* \in \mathcal{D}$ be such that

$$F(\pi^*) + I(\pi^*) \leq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\} + \varepsilon.$$

Note in particular that $I(\pi^*) < \infty$. Next let $\Theta \in \mathcal{P}_\infty^1$ be such that $\varpi_\Theta = \pi^*$ and $\mathcal{J}(\Theta) \leq I(\pi^*) + \varepsilon$. Recall the canonical variables $(\tilde{\mathbf{b}}, \mathbf{b}, \boldsymbol{\rho}, \tilde{\boldsymbol{\sigma}})$ introduced in (3.1). We claim that without loss of generality one can assume that

$$\boldsymbol{\rho}(A \times [0, t]) = \int_0^t \delta_A(\boldsymbol{\psi}(s)) ds, \quad t \in [0, T], A \in \mathcal{B}(\mathbb{R}^d), \text{ a.s. } \Theta, \quad (7.1)$$

where $\boldsymbol{\psi}$ is an \mathbb{R}^d -valued \mathcal{G}_t -progressively measurable process such that

$$\mathbf{E}^\Theta \int_0^T \|\boldsymbol{\psi}(s)\|^2 ds < \infty.$$

To see this, define $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{\rho}}$ on Ξ as follows. For $t \in [0, T]$, $A \in \mathcal{B}(\mathbb{R}^d)$ and $\xi = (\tilde{b}, b, \rho, \tilde{\sigma}) \in \Xi$,

$$\boldsymbol{\psi}[\xi](t) \doteq \int_{\mathbb{R}^d} y \rho(dy|t), \quad \tilde{\boldsymbol{\rho}}[\xi](A \times [0, t]) \doteq \int_0^t \delta_A(\boldsymbol{\psi}[\xi](s)) ds,$$

where $\rho(dy dt) = \rho(dy|t) dt$. Then

$$\int_{\mathbb{R}^d \times [0, t]} y \tilde{\boldsymbol{\rho}}(dy ds) = \int_0^t \boldsymbol{\psi}(s) ds = \int_{\mathbb{R}^d \times [0, t]} y \boldsymbol{\rho}(dy ds) \quad (7.2)$$

and

$$\int_{\mathbb{R}^d \times [0, t]} \|y\|^2 \tilde{\boldsymbol{\rho}}(dy ds) = \int_0^t \|\boldsymbol{\psi}(s)\|^2 ds \leq \int_{\mathbb{R}^d \times [0, t]} \|y\|^2 \boldsymbol{\rho}(dy ds). \quad (7.3)$$

Thus with $\tilde{\Theta} \doteq \Theta \circ (\tilde{\mathbf{b}}, \mathbf{b}, \tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{\sigma}})^{-1}$, it follows from (7.2) and (7.3) that $\tilde{\Theta} \in \mathcal{P}_\infty^1$, $\varpi_{\tilde{\Theta}} = \varpi_\Theta = \pi^*$ and $\mathcal{J}(\tilde{\Theta}) \leq \mathcal{J}(\Theta) \leq I(\pi^*) + \varepsilon$. This proves the claim. Henceforth we assume (7.1).

We now construct a filtered probability space $\Upsilon \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}_{0 \leq t \leq T}, \bar{\mathbf{P}})$ as follows. Let $\bar{\Omega} \doteq \Xi^{\otimes \infty}$, $\bar{\mathcal{F}}$ be the $\bar{\mathbf{P}}$ -completion of $\mathcal{B}(\Xi^{\otimes \infty})$, where $\bar{\mathbf{P}} \doteq \Theta^{\otimes \infty}$. Define canonical variables on this space as follows: For $\xi \doteq (\xi_i)_{i \in \mathbb{N}} \in \bar{\Omega}$, where $\xi_i \doteq (\tilde{b}_i, b_i, \rho_i, \tilde{\sigma}_i)$

$$\tilde{\mathbf{b}}_i(\xi) \doteq \tilde{b}_i, \mathbf{b}_i(\xi) \doteq b_i, \boldsymbol{\rho}_i(\xi) \doteq \rho_i, \tilde{\boldsymbol{\sigma}}_i(\xi) = \tilde{\sigma}_i, i \in \mathbb{N}.$$

Let $\hat{\mathcal{F}}_t$ be the $\bar{\mathbf{P}}$ -completion of $\sigma\{\tilde{\mathbf{b}}_i(s), \mathbf{b}_i(s), \boldsymbol{\rho}_i(A \times [0, s]), \tilde{\boldsymbol{\sigma}}_i, i \in \mathbb{N}, s \leq t, A \in \mathcal{B}(\mathbb{R}^d)\}$. Define

$$\boldsymbol{\psi}_i(t) \doteq \int_{\mathbb{R}^d} y \boldsymbol{\rho}_i(dy|t), i \in \mathbb{N}, t \in [0, T].$$

Note that by (7.1), a.s. $\bar{\mathbf{P}}$, $\boldsymbol{\rho}_i(A \times [0, t]) = \int_0^t \delta_A(\boldsymbol{\psi}_i(s)) ds$ for $A \in \mathcal{B}(\mathbb{R}^d)$, $t \in [0, T]$ and $i \in \mathbb{N}$. With above choices of Υ and canonical random variables, let $(\boldsymbol{\nu}_i^n)_{i=1}^n$ and $\tilde{\mu}^n$ be defined as above Lemma 5.1, with $\Pi^n = \Theta_{(4)}^{\otimes n}$ and $(\mathbf{S}^n, \tilde{\boldsymbol{\beta}}^n, \boldsymbol{\beta}^n)$ replaced with $(\tilde{\boldsymbol{\sigma}}^n, \tilde{\mathbf{b}}^n, \mathbf{b}^n)$, where $\tilde{\boldsymbol{\sigma}}^n \doteq (\tilde{\boldsymbol{\sigma}}_i)_{i=1}^n$, $\tilde{\mathbf{b}}^n \doteq (\tilde{b}_i)_{i=1}^n$ and $\mathbf{b}^n \doteq (b_i)_{i=1}^n$. Note in particular that $\boldsymbol{\nu}_i^n = \Theta_{(4)}$ for each $i = 1, \dots, n$. It follows from Lemma 5.1 that

$$-\frac{1}{n} \log \mathbf{E} e^{-nF(\mu^n)} \leq \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\boldsymbol{\nu}_i^n \| \theta) + \frac{1}{2} \int_0^T \|\boldsymbol{\psi}_i(s)\|^2 ds \right) + F(\tilde{\mu}^n) \right]. \quad (7.4)$$

Note that

$$\begin{aligned} & \bar{\mathbf{E}} \left[\frac{1}{n} \sum_{i=1}^n \left(R(\boldsymbol{\nu}_i^n \| \theta) + \frac{1}{2} \int_0^T \|\boldsymbol{\psi}_i(s)\|^2 ds \right) \right] \\ &= R(\Theta_{(4)} \| \theta) + \frac{1}{2n} \sum_{i=1}^n \bar{\mathbf{E}} \left[\int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}_i(dy ds) \right] \\ &= R(\Theta_{(4)} \| \theta) + \mathbf{E}^\Theta \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy ds) \right] \\ &= \mathcal{J}(\Theta), \end{aligned} \quad (7.5)$$

where the second equality holds since $\boldsymbol{\rho}_i$ are i.i.d. under $\bar{\mathbf{P}}$. The above calculation shows that (5.5) holds with C_0 replaced by $\mathcal{J}(\Theta) < \infty$. Define (Q^n, ν^n) as in (5.6) and (5.7) (with $(\mathbf{S}^n, \tilde{\boldsymbol{\beta}}^n, \boldsymbol{\beta}^n)$ replaced by $(\tilde{\boldsymbol{\sigma}}^n, \tilde{\mathbf{b}}^n, \mathbf{b}^n)$ and $(\boldsymbol{\rho}_i^n)_{i=1}^n$ replaced by $(\boldsymbol{\rho}_i)_{i=1}^n$). It follows from Lemmas 5.2 and 5.4 that $\{(Q^n, \nu^n, \tilde{\mu}^n)\}$ is a tight sequence of $\mathcal{P}(\Xi) \times \mathcal{P}(\mathbb{R}^d) \times \mathcal{D}$ -valued random variables. Suppose $(Q^*, \nu^*, \tilde{\mu}^*)$ is a weak limit point of the sequence given on some probability space $(\Omega^*, \mathcal{F}^*, \mathbf{P}^*)$. From Lemmas 5.3 and 5.5 it follows that $Q^* \in \mathcal{P}_\infty$, $Q_{(4)}^* = \nu^*$ and $\tilde{\mu}^* = \varpi_{Q^*}$ a.s. \mathbf{P}^* . By law of large numbers $Q^* = \Theta$ a.s. \mathbf{P}^* and hence we have $\tilde{\mu}^* = \varpi_\Theta = \pi^*$ a.s. \mathbf{P}^* . Combining above observations with (7.4) and (7.5) we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{E} e^{-nF(\mu^n)} &\leq \mathcal{J}(\Theta) + F(\pi^*) \\ &\leq I(\pi^*) + F(\pi^*) + \varepsilon \\ &\leq \inf_{\pi \in \mathcal{D}} \{F(\pi) + I(\pi)\} + 2\varepsilon, \end{aligned}$$

where the last two inequalities follow from the choices of Θ and π^* . Since $\varepsilon > 0$ is arbitrary, the desired Laplace lower bound follows. \square

8. I is a rate function.

In this section we will prove that the function I defined in (3.4) is a rate function, namely we prove the following.

Proposition 8.1. *For each fixed $M < \infty$, the set $F_M \doteq \{\pi \in \mathcal{D} : I(\pi) \leq M\}$ is compact.*

Proof. Let $\{\pi_n\}_{n \in \mathbb{N}} \subset F_M$. Then for each $n \in \mathbb{N}$ there exists $\Theta^n \in \mathcal{P}_\infty^1$ such that $\varpi_{\Theta^n} = \pi_n$ and $\mathcal{J}(\Theta^n) \leq M + \frac{1}{n}$. In particular

$$\sup_{n \in \mathbb{N}} \left(R(\Theta_{(4)}^n \| \theta) + \mathbf{E}^{\Theta^n} \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right] \right) \leq M + 1.$$

Using the above bound we now argue that $\{\Theta^n\}$ is tight. The proof is similar to that of Lemma 5.2 and so we only provide a sketch. Note that $\Theta_{(2)}^n$ is the d -dimensional Wiener measure for each n and so the tightness of $\{\Theta_{(2)}^n\}$ is immediate. Next, recall tightness functions G_j for $j = 1, 2, 3$ and the function g_3 in the proof of Lemma 5.2. Since $G_1(\Theta_{(3)}^n) = \mathbf{E}^{\Theta^n} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \leq 2(M + 1)$, $\{\Theta_{(3)}^n\}$ is tight. To argue tightness of $\{\Theta_{(1)}^n\}$, define $\mathbf{u} : \Xi \rightarrow \mathcal{C}$ as $\mathbf{u}(t)[\xi] \doteq \int_{\mathbb{R}^d \times [0, t]} y \rho(dy ds)$ for $t \in [0, T]$. It follows from Cauchy–Schwarz inequality that $G_2(\Theta^n \circ \mathbf{u}^{-1}) = \mathbf{E}^{\Theta^n} \int_0^T \|\dot{\mathbf{u}}(s)\|^2 ds \leq \mathbf{E}^{\Theta^n} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \leq 2(M + 1)$. So $\{\Theta^n \circ \mathbf{u}^{-1}\}$ is tight and consequently $\{\Theta^n \circ (\mathbf{b}, \mathbf{u})^{-1}\}$ is also tight. This implies tightness of $\{\Theta_{(1)}^n\}$ since $\Theta_{(1)}^n = \Theta^n \circ (g_3(\mathbf{b}, \mathbf{u}))^{-1}$ and g_3 is continuous. Finally, tightness of $\{\Theta_{(4)}^n\}$ follows since $G_3(\Theta_{(4)}^n) = R(\Theta_{(4)}^n \| \theta) \leq M + 1$. This proves the tightness of $\{\Theta^n\}$.

Now suppose Θ^n converges along a subsequence (labeled as $\{n\}$ for simplicity) to $\Theta \in \mathcal{P}(\Xi)$. Using arguments similar to those in the proof of Lemma 5.3, we now show that $\Theta \in \mathcal{P}_\infty$ and $\mathcal{J}(\Theta) < \infty$. We need to show that $\Theta_{(4)} \ll \theta$ and verify properties (1)–(3) of Section 3.2 for Θ . It is clear that (1) holds. Using the lower semi-continuity of the function $R(\cdot \| \theta)$ and the map in (5.10), we have

$$\begin{aligned} \mathcal{J}(\Theta) &= R(\Theta_{(4)} \| \theta) + \mathbf{E}^\Theta \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right] \\ &\leq \liminf_{n \rightarrow \infty} \left(R(\Theta_{(4)}^n \| \theta) + \mathbf{E}^{\Theta^n} \left[\frac{1}{2} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \rho(dy ds) \right] \right) \leq M. \end{aligned} \quad (8.1)$$

This implies (2), the property $\Theta_{(4)} \ll \theta$, and that $\mathcal{J}(\Theta) < \infty$. Finally, recalling that the map (5.12) is bounded and continuous, we have

$$\begin{aligned} &\mathbf{E}^\Theta \left(\left| \tilde{\mathbf{b}}_t - \mathbf{b}_t - \int_{\mathbb{R}^d \times [0, t]} y \rho(dy ds) \right| \wedge 1 \right) \\ &= \lim_{n \rightarrow \infty} \mathbf{E}^{\Theta^n} \left(\left| \tilde{\mathbf{b}}_t - \mathbf{b}_t - \int_{\mathbb{R}^d \times [0, t]} y \rho(dy ds) \right| \wedge 1 \right) = 0 \end{aligned}$$

for each $t \in [0, T]$, which verifies (3). Thus we have shown that $\Theta \in \mathcal{P}_\infty$ and $\mathcal{J}(\Theta) < \infty$.

We will now use arguments similar to those in Lemma 5.4 to show that $\{\varpi_{\Theta^n}\}$ is pre-compact in $\mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. The tightness of $\{\varpi_{\Theta^n}(t)\}$ in $\mathcal{M}(\mathbb{R}^d)$ for each $t \in [0, T]$ follows from the estimate

$$G_4(\varpi_{\Theta^n}(t)) \leq 2\mathbf{E}^{\Theta^n} [\|\mathbf{b}_t\|^2] + 2T\mathbf{E}^{\Theta^n} \left[\int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy ds) \right] \leq 2dT + 4T(M+1). \quad (8.2)$$

Next we proceed to consider fluctuations. Fix $\delta > 0$ and consider $0 \leq t_1 \leq t_2 \leq T$ with $|t_1 - t_2| < \delta$. As in the proof of Lemma 5.4 we can estimate

$$\begin{aligned} d_{BL}(\varpi_{\Theta^n}(t_1), \varpi_{\Theta^n}(t_2)) &\leq \sup_{\|f\|_{BL} \leq 1} \mathbf{E}^{\Theta^n} \left| f(\tilde{\mathbf{b}}(t_1)) - f(\tilde{\mathbf{b}}(t_2)) \right| \mathbf{1}_{\{\tilde{\sigma} > \int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds\}} \\ &\quad + \sup_{\|f\|_{BL} \leq 1} \mathbf{E}^{\Theta^n} \left| f(\tilde{\mathbf{b}}(t_2)) \left(\mathbf{1}_{\{\tilde{\sigma} > \int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds\}} - \mathbf{1}_{\{\tilde{\sigma} > \int_0^{t_2} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds\}} \right) \right| \\ &\doteq \bar{\mathcal{T}}_1^n + \bar{\mathcal{T}}_2^n. \end{aligned}$$

For $\bar{\mathcal{T}}_1^n$, using Cauchy–Schwarz inequality as in Lemma 5.4

$$\begin{aligned} \sup_{|t_1 - t_2| < \delta} \bar{\mathcal{T}}_1^n &\leq \sup_{|t_1 - t_2| < \delta} \left(\mathbf{E}^{\Theta^n} \|\mathbf{b}(t_1) - \mathbf{b}(t_2)\|^2 \right)^{\frac{1}{2}} + \sqrt{\delta} \left(\mathbf{E}^{\Theta^n} \int_{\mathbb{R}^d \times [0, T]} \|y\|^2 \boldsymbol{\rho}(dy ds) \right)^{\frac{1}{2}} \\ &\leq \sqrt{d\delta} + \sqrt{2(M+1)\delta} \rightarrow 0, \end{aligned}$$

uniformly in n , as $\delta \rightarrow 0$.

Consider now $\bar{\mathcal{T}}_2^n$. It follows from (2.2) and (8.2) that

$$\langle \zeta, \varpi_{\Theta^n}(t) \rangle \leq \tilde{C}_\zeta \mathbf{E}^{\Theta^n} (1 + \|\tilde{\mathbf{b}}_t\|^2) \leq \tilde{C}_\zeta (1 + 2dT + 4T(M+1)) \doteq \bar{C}_0.$$

So we have

$$\left| \int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds - \int_0^{t_2} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds \right| \leq \bar{C}_0 \delta$$

and hence

$$\begin{aligned} \sup_{|t_1 - t_2| < \delta} \bar{\mathcal{T}}_2^n &\leq \sup_{|t_1 - t_2| < \delta} \Theta^n \left(\int_0^{t_1} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds < \tilde{\sigma} \leq \int_0^{t_2} \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds \right) \\ &\leq \max_{j \in \mathbb{N}_0} \Theta_{(4)}^n([j\bar{C}_0\delta, (j+2)\bar{C}_0\delta]). \end{aligned}$$

Since $\Theta^n \rightarrow \Theta \in \mathcal{P}_\infty$, it then follows from (5.20) that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sup_{|t_1 - t_2| < \delta} \bar{\mathcal{T}}_2^n &\leq \limsup_{n \rightarrow \infty} \max_{j \in \mathbb{N}_0} \Theta_{(4)}^n([j\bar{C}_0\delta, (j+2)\bar{C}_0\delta]) \\ &= \max_{j \in \mathbb{N}_0} \Theta_{(4)}([j\bar{C}_0\delta, (j+2)\bar{C}_0\delta]) \rightarrow 0, \end{aligned}$$

as $\delta \rightarrow 0$, where the convergence to 0 is argued using an estimate as in (5.21). Combining above two convergence results of \bar{T}_1^n and \bar{T}_2^n gives us pre-compactness of $\{\varpi_{\Theta^n}\}$ in $\mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$.

Next, let ϖ_{Θ^n} converge to π^* along a further subsequence (labeled once again as $\{n\}$). Then we have $(\Theta^n, \varpi_{\Theta^n}) \rightarrow (\Theta, \pi^*)$ in $\mathcal{P}(\Xi) \times \mathbb{C}([0, T] : \mathcal{M}(\mathbb{R}^d))$. Fix $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t \in [0, T]$. Recall the function g_t associated with f defined in (5.24). Note that, for $t \in [0, T]$,

$$\langle f, \varpi_{\Theta^n}(t) \rangle = g_t(\Theta^n, \varpi_{\Theta^n}), \quad \lim_{n \rightarrow \infty} \langle f, \varpi_{\Theta^n}(t) \rangle = \langle f, \pi^*(t) \rangle.$$

We claim that $g_t(\Theta^n, \varpi_{\Theta^n}) \rightarrow g_t(\Theta, \pi^*)$ as $n \rightarrow \infty$ for all $t \in [0, T]$. Once the claim is verified, we will have $\langle f, \pi^*(t) \rangle = g_t(\Theta, \pi^*)$ and consequently by Lemma 3.1 $\pi^* = \varpi_{\Theta}$. This along with (8.1) will say that $I(\pi^*) = I(\varpi_{\Theta}) \leq \mathcal{J}(\Theta) \leq M$. Hence $\pi^n = \varpi_{\Theta^n} \rightarrow \varpi_{\Theta} = \pi^* \in F_M$ completing the proof of the Proposition.

To see the claim that $g_t(\Theta^n, \varpi_{\Theta^n}) \rightarrow g_t(\Theta, \pi^*)$ as $n \rightarrow \infty$, note that

$$|g_t(\Theta^n, \varpi_{\Theta^n}) - g_t(\Theta, \pi^*)| \leq |g_t(\Theta^n, \varpi_{\Theta^n}) - g_t(\Theta^n, \pi^*)| + |g_t(\Theta^n, \pi^*) - g_t(\Theta, \pi^*)|. \quad (8.3)$$

Since $\Theta^n \rightarrow \Theta \in \mathcal{P}_{\infty}$, by Remark 3.1 we have

$$\lim_{n \rightarrow \infty} |g_t(\Theta^n, \pi^*) - g_t(\Theta, \pi^*)| = 0. \quad (8.4)$$

For the first term on the right hand side of (8.3), we have

$$\begin{aligned} & |g_t(\Theta^n, \varpi_{\Theta^n}) - g_t(\Theta^n, \pi^*)| \\ & \leq \mathbf{E}^{\Theta^n} \left| f(\tilde{\mathbf{b}}_t) \left(\mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds\}} - \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \pi^*(s) \rangle ds\}} \right) \right| \\ & \leq \|f\|_{\infty} \Theta^n \left(\tilde{\sigma} \text{ lies between } \int_0^t \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds \text{ and } \int_0^t \langle \zeta, \pi^*(s) \rangle ds \right). \end{aligned}$$

Similar to the proof of (5.23), with ζ_K defined as above (5.27), one can show using (5.28) and (8.2) that for $t \in [0, T]$,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left| \int_0^t \langle \zeta_K, \varpi_{\Theta^n}(s) \rangle ds - \int_0^t \langle \zeta_K, \pi^*(s) \rangle ds \right| = 0, \\ & \lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \left| \int_0^t \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds - \int_0^t \langle \zeta_K, \varpi_{\Theta^n}(s) \rangle ds \right| = 0, \\ & \lim_{K \rightarrow \infty} \left| \int_0^t \langle \zeta, \pi^*(s) \rangle ds - \int_0^t \langle \zeta_K, \pi^*(s) \rangle ds \right| = 0. \end{aligned}$$

Hence as $n \rightarrow \infty$,

$$\left| \int_0^t \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds - \int_0^t \langle \zeta, \pi^*(s) \rangle ds \right| \rightarrow 0.$$

So for any $\delta > 0$, using (5.18) and (5.20)

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} |g_t(\Theta^n, \varpi_{\Theta^n}) - g_t(\Theta^n, \pi^*)| \\
& \leq \|f\|_\infty \limsup_{n \rightarrow \infty} \Theta^n \left(\tilde{\sigma} \text{ lies between } \int_0^t \langle \zeta, \varpi_{\Theta^n}(s) \rangle ds \text{ and } \int_0^t \langle \zeta, \pi^*(s) \rangle ds \right) \\
& \leq \|f\|_\infty \limsup_{n \rightarrow \infty} \max_{j \in \mathbb{N}_0} \Theta_{(4)}^n([j\delta, (j+1)\delta]) \\
& = \|f\|_\infty \max_{j \in \mathbb{N}_0} \Theta_{(4)}([j\delta, (j+1)\delta]).
\end{aligned}$$

Letting $\delta \rightarrow 0$ in the above display and using a similar estimate as in (5.21) we have that $\limsup_{n \rightarrow \infty} |g_t(\Theta^n, \varpi_{\Theta^n}) - g_t(\Theta^n, \pi^*)| = 0$. This together with (8.3) and (8.4) proves the claim that for every $t \in [0, T]$, $g_t(\Theta^n, \varpi_{\Theta^n}) \rightarrow g_t(\Theta, \pi^*)$ as $n \rightarrow \infty$. \square

9. Proof of Proposition 4.1

Our starting point for the proof of Proposition 4.1 will be the representation in (4.5). We first consider the simpler setting where for fixed $\Pi \in \mathcal{P}(\mathbb{S})$ and a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ as in Section 4, the filtration $\bar{\mathcal{F}}_t$ is replaced by $\bar{\mathcal{F}}_t$ where recall that $\bar{\mathcal{F}}_t \doteq \sigma\{\bar{\mathcal{F}}_t^{\bar{W}, \bar{X}} \cup \bar{\mathcal{N}}\}$, $\bar{\mathcal{N}}$ is the collection of all $\bar{\mathbf{P}}$ -null sets, and $\bar{\mathcal{F}}_t^{\bar{W}, \bar{X}} \doteq \sigma\{\bar{X}, \bar{W}(s) : s \leq t\}$. We write $\bar{\Upsilon}_\Pi \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{\mathbf{P}})$ and consider the collection of processes

$$\begin{aligned}
& \mathcal{A}(\bar{\Upsilon}_\Pi) \doteq \{\bar{\psi} : \text{the process } \bar{\psi} \text{ is } \bar{\mathcal{F}}_t\text{-progressively measurable} \\
& \text{and } \bar{\mathbf{E}} \int_0^T \|\bar{\psi}(s)\|^2 ds < \infty\}.
\end{aligned}$$

For each $N < \infty$, let

$$\mathcal{A}^N(\bar{\Upsilon}_\Pi) \doteq \left\{ \bar{\psi} \in \mathcal{A}(\bar{\Upsilon}_\Pi) : \int_0^T \|\bar{\psi}(s)\|^2 ds \leq N, \bar{\mathbf{P}}\text{-a.s.} \right\}. \quad (9.1)$$

For $\mathcal{A} \subset \mathcal{A}(\bar{\Upsilon}_\Pi)$ and $g \in \mathbb{M}_b(\mathcal{C} \times \mathbb{S})$, define

$$\Lambda(\mathcal{A}, g) \doteq \inf_{\bar{\psi} \in \mathcal{A}} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + g \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right]. \quad (9.2)$$

The following lemma gives a convenient approximation result. The proof is similar to that of [3, Lemma 3.4] (see also proof of [2, Theorem 3.1]), but we provide details for completeness. Recall from Section 4 that $\tilde{\mathbf{P}}$ is the Wiener measure on $(\tilde{\Omega}, \tilde{\mathcal{F}}) = (\mathcal{C}, \mathcal{B}(\mathcal{C}))$.

Lemma 9.1. *Fix $\Pi \in \mathcal{P}(\mathbb{S})$. Let $\{f_n\}$ be a uniformly bounded sequence of real-valued measurable functions on $\mathcal{C} \times \mathbb{S}$ converging to f a.s. $\tilde{\mathbf{P}} \times \Pi$. Then for every $N < \infty$, $\Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f)$ as $n \rightarrow \infty$.*

Proof. Fix $\varepsilon > 0$. For each $n \in \mathbb{N}$ we pick an element $\bar{\psi}_{n,\varepsilon} \in \mathcal{A}^N(\bar{\Upsilon}_\Pi)$ such that

$$\Lambda(\mathcal{A}^N, f_n) \geq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_{n,\varepsilon}(s)\|^2 ds + f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right) \right] - \varepsilon.$$

By definition, for each $n \in \mathbb{N}$

$$\bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_{n,\varepsilon}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right) \right] \geq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f).$$

We claim that as $n \rightarrow \infty$,

$$\bar{\mathbf{E}} f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right) - \bar{\mathbf{E}} f \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right) \rightarrow 0.$$

Suppose for now that the claim holds. Combining above three displays we have $\liminf_{n \rightarrow \infty} \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f_n) \geq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f) - \varepsilon$. Since $\varepsilon > 0$ is arbitrary, this shows $\liminf_{n \rightarrow \infty} \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f_n) \geq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f)$. Now in order to prove the claim, from [2, Lemma 2.8(b)], it suffices to show that the relative entropies

$$R \left(\bar{\mathbf{P}} \circ \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right)^{-1} \parallel \tilde{\mathbf{P}} \times \Pi \right)$$

are uniformly bounded in n . For this, first consider the probability measure $\bar{\mathbf{P}}^{n,\varepsilon}$ defined by

$$\frac{d\bar{\mathbf{P}}^{n,\varepsilon}}{d\bar{\mathbf{P}}} \doteq \exp \left\{ - \int_0^T \bar{\psi}_{n,\varepsilon}(s) d\bar{W}(s) - \frac{1}{2} \int_0^T \|\bar{\psi}_{n,\varepsilon}(s)\|^2 ds \right\}.$$

By Girsanov's theorem, on the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}}^{n,\varepsilon})$, $\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds$ is an $\{\bar{\mathcal{F}}_t\}$ -Brownian motion independent of the $\bar{\mathcal{F}}_0$ -measurable random variable \bar{X} , and $\bar{\mathbf{P}}^{n,\varepsilon} \circ (\bar{X})^{-1} = \Pi$. So we have

$$\bar{\mathbf{P}}^{n,\varepsilon} \circ \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right)^{-1} = \tilde{\mathbf{P}} \times \Pi$$

and hence, for every $n \in \mathbb{N}$,

$$\begin{aligned} R \left(\bar{\mathbf{P}} \circ \left(\bar{W} + \int_0^\cdot \bar{\psi}_{n,\varepsilon}(s) ds, \bar{X} \right)^{-1} \parallel \tilde{\mathbf{P}} \times \Pi \right) \\ \leq R(\bar{\mathbf{P}} \parallel \bar{\mathbf{P}}^{n,\varepsilon}) = \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_{n,\varepsilon}(s)\|^2 ds \right] \leq \frac{N}{2}. \end{aligned}$$

Thus the claim holds and we have $\liminf_{n \rightarrow \infty} \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f_n) \geq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f)$.

We now prove $\limsup_{n \rightarrow \infty} \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f_n) \leq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f)$. Pick an element $\bar{\psi}_\varepsilon \in \mathcal{A}^N(\bar{\Upsilon}_\Pi)$ such that

$$\bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X} \right) \right] \leq \Lambda(\mathcal{A}^N(\bar{\Upsilon}_\Pi), f) + \varepsilon.$$

Since $\bar{\mathbf{P}} \circ (\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X})^{-1} \ll \tilde{\mathbf{P}} \times \Pi$, we have that as $n \rightarrow \infty$,

$$\bar{\mathbf{E}} f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X} \right) \rightarrow \bar{\mathbf{E}} f \left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X} \right).$$

Also, by definition, for each $n \in \mathbb{N}$

$$\Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f_n) \leq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 ds + f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X} \right) \right].$$

Combining the last three displays gives $\limsup_{n \rightarrow \infty} \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f_n) \leq \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f) + \varepsilon$. The result follows since $\varepsilon > 0$ is arbitrary. \square

Remark 9.1. From the proof above, we see that Lemma 9.1 remains valid even if $\{\bar{\mathcal{F}}_t\}$ is replaced by the general $\{\hat{\mathcal{F}}_t\}$ as long as \bar{W} is a Brownian motion with respect to this filtration and \bar{X} is measurable with respect to $\hat{\mathcal{F}}_0$.

The following lemma proves Proposition 4.1 in the special case where the general filtration $\hat{\mathcal{F}}_t$ is replaced by $\bar{\mathcal{F}}_t$.

Lemma 9.2. *Let $f \in \mathbb{M}_b(\mathcal{C} \times \mathbb{S})$. Then*

$$\begin{aligned} & -\log \mathbf{E} [\exp(-f(B, X))] \\ &= \inf_{\Pi, \tilde{\Upsilon}_\Pi} \inf_{\bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi)} \left\{ R(\Pi \| \rho) + \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] \right\}, \end{aligned} \quad (9.3)$$

where the outer infimum is over all $\Pi \in \mathcal{P}(\mathbb{S})$ and all systems $\tilde{\Upsilon}_\Pi$ of the form introduced at the beginning of the section.

Proof. In view of (4.3), it suffices to prove that for every $\Pi \in \mathcal{P}(\mathbb{S})$ such that $R(\Pi \| \rho) < \infty$,

$$\int_{\mathbb{S}} \tilde{f}(x) \Pi(dx) = \inf_{\bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi)} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right], \quad (9.4)$$

where \tilde{f} is defined in (4.2). We claim that it suffices to prove (9.4) for $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. Assume for now that the claim holds and so we assume that $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. We first argue that LHS \leq RHS in (9.4). Fix $\bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi)$. Since $\bar{\psi}$ is $\bar{\mathcal{F}}_t$ -progressively measurable, there exists (cf. [22, Exercise 1.5.6]) a $(\mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{S}) \times \mathcal{B}(\mathcal{C})) / \mathcal{B}(\mathbb{R}^d)$ -measurable map $F: [0, T] \times \mathbb{S} \times \mathcal{C} \rightarrow \mathbb{R}^d$ such that

$$\bar{\psi}(s, \bar{\omega}) = F(s, \bar{X}(\bar{\omega}), \bar{W}_{s \wedge \cdot}(\bar{\omega})), \quad \bar{\mathbf{P}}\text{-a.s. } (s, \bar{\omega}) \in [0, T] \times \bar{\Omega}.$$

Recall the Wiener space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ introduced in Section 4. Define a collection of processes $\{\tilde{\psi}_x\}_{x \in \mathbb{S}}$ on $\tilde{\Omega}$ as

$$\tilde{\psi}_x(s, \tilde{\omega}) \doteq F(s, x, \tilde{W}_{s \wedge \cdot}(\tilde{\omega})), \quad (x, s, \tilde{\omega}) \in \mathbb{S} \times [0, T] \times \tilde{\Omega}.$$

Then $\tilde{\psi}_x \in \tilde{\mathcal{A}}$ for Π -a.e. $x \in \mathbb{S}$ where $\tilde{\mathcal{A}}$ was defined below (4.3). By independence between \bar{X}

and \bar{W} ,

$$\begin{aligned}
& \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] \\
&= \int_{\mathbb{S}} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|F(s, x, \bar{W}_{s\wedge\cdot})\|^2 ds + f \left(\bar{W} + \int_0^\cdot F(s, x, \bar{W}_{s\wedge\cdot}) ds, x \right) \right] \Pi(dx) \\
&= \int_{\mathbb{S}} \tilde{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\tilde{\psi}_x(s)\|^2 ds + f \left(\tilde{W} + \int_0^\cdot \tilde{\psi}_x(s) ds, x \right) \right] \Pi(dx) \\
&\geq \int_{\mathbb{S}} \tilde{f}(x) \Pi(dx),
\end{aligned}$$

where the last inequality is from (4.4). Taking the infimum over all $\bar{\psi} \in \mathcal{A}(\bar{\Upsilon}_\Pi)$ in the above inequality implies that LHS \leq RHS in (9.4).

We now consider the reverse inequality. Fix $\varepsilon \in (0, 1)$. For each $x \in \mathbb{S}$, let $\tilde{\psi}_{x,\varepsilon} \in \tilde{\mathcal{A}}$ be a control in (4.4) such that

$$\tilde{f}(x) \geq \tilde{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\tilde{\psi}_{x,\varepsilon}(s)\|^2 ds + f \left(\tilde{W} + \int_0^\cdot \tilde{\psi}_{x,\varepsilon}(s) ds, x \right) \right] - \varepsilon. \quad (9.5)$$

We will now carefully select a countable sub-collection from $\{\tilde{\psi}_{x,\varepsilon}\}_{x \in \mathbb{S}}$ and use it to construct an $\tilde{\mathcal{F}}_t$ -progressively measurable process $\bar{\psi}_\varepsilon \in \mathcal{A}(\bar{\Upsilon}_\Pi)$. From (9.5) we have

$$\sup_{x \in \mathbb{S}} \tilde{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\tilde{\psi}_{x,\varepsilon}(s)\|^2 ds \right] \leq 2\|f\|_\infty + 1.$$

Using this it is easy to check that $\{\int_0^\cdot \tilde{\psi}_{x,\varepsilon}(s) ds\}_{x \in \mathbb{S}}$ is tight in \mathcal{C} and thus so is $\{\tilde{W}^{x,\varepsilon} \doteq \tilde{W} + \int_0^\cdot \tilde{\psi}_{x,\varepsilon}(s) ds\}_{x \in \mathbb{S}}$. Then there exists a compact $K_\varepsilon \subset \mathcal{C}$ such that

$$\sup_{x \in \mathbb{S}} \tilde{\mathbf{P}} \left(\tilde{W}^{x,\varepsilon} \notin K_\varepsilon \right) \leq \varepsilon / \|f\|_\infty. \quad (9.6)$$

Let \tilde{K}_ε be a compact subset of \mathbb{S} such that

$$\Pi(\mathbb{S} \setminus \tilde{K}_\varepsilon) \leq \varepsilon / \|f\|_\infty.$$

Since f is continuous, so is \tilde{f} . In particular, f is uniformly continuous on $K_\varepsilon \times \tilde{K}_\varepsilon$ and \tilde{f} is uniformly continuous on \tilde{K}_ε . So there exists some $M_\varepsilon \in \mathbb{N}$ and a finite partition $\{B_i : i = 1, \dots, M_\varepsilon\}$ of \tilde{K}_ε such that

$$\max_{i=1, \dots, M_\varepsilon} \sup_{\phi \in K_\varepsilon, x_1, x_2 \in B_i} |f(\phi, x_1) - f(\phi, x_2)| + |\tilde{f}(x_1) - \tilde{f}(x_2)| \leq \varepsilon. \quad (9.7)$$

Now for each $i = 1, \dots, M_\varepsilon$, fix a $y_i \in B_i$ and define

$$\bar{\psi}_\varepsilon(s, \bar{\omega}) \doteq \begin{cases} \tilde{\psi}_{y_i, \varepsilon}(s, \bar{W}(\bar{\omega})), & \bar{\omega} \in \bar{X}^{-1}(B_i), \\ 0, & \bar{\omega} \notin \bar{X}^{-1}(\tilde{K}_\varepsilon). \end{cases}$$

Since $\tilde{\psi}_{y_i, \varepsilon}$ is $\tilde{\mathcal{F}}_t$ -progressively measurable for each $i = 1, \dots, M_\varepsilon$ and \bar{X} is $\bar{\mathcal{F}}_0$ -measurable, $\bar{\psi}_\varepsilon$ is $\tilde{\mathcal{F}}_t$ -progressively measurable. For $x \in B_i$, $i \in \{1, \dots, M_\varepsilon\}$, we have

$$\begin{aligned} \bar{\mathbf{E}}f\left(\bar{W} + \int_0^\cdot \tilde{\psi}_{y_i, \varepsilon}(s, \bar{W}) ds, x\right) &= \tilde{\mathbf{E}}f\left(\tilde{W}^{y_i, \varepsilon}, x\right) \\ &\leq \tilde{\mathbf{E}}\left[f\left(\tilde{W}^{y_i, \varepsilon}, x\right) \mathbf{1}_{\{\tilde{W}^{y_i, \varepsilon} \in K_\varepsilon\}}\right] + \varepsilon \\ &\leq \tilde{\mathbf{E}}\left[f\left(\tilde{W}^{y_i, \varepsilon}, y_i\right) \mathbf{1}_{\{\tilde{W}^{y_i, \varepsilon} \in K_\varepsilon\}}\right] + 2\varepsilon \\ &\leq \tilde{\mathbf{E}}f\left(\tilde{W}^{y_i, \varepsilon}, y_i\right) + 3\varepsilon, \end{aligned}$$

where the first and third inequalities use (9.6), and the second inequality follows from (9.7). Using this, the definition of \tilde{K}_ε , and the independence between \bar{W} and \bar{X} , we have

$$\begin{aligned} &\bar{\mathbf{E}}f\left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X}\right) \\ &= \sum_{i=1}^{M_\varepsilon} \bar{\mathbf{E}}\left[\mathbf{1}_{\{\bar{X} \in B_i\}} f\left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X}\right)\right] + \bar{\mathbf{E}}\left[\mathbf{1}_{\{\bar{X} \notin \tilde{K}_\varepsilon\}} f\left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X}\right)\right] \\ &\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \bar{\mathbf{E}}f\left(\bar{W} + \int_0^\cdot \tilde{\psi}_{y_i, \varepsilon}(s, \bar{W}) ds, x\right) \Pi(dx) + \varepsilon \\ &\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{\mathbf{E}}f\left(\tilde{W}^{y_i, \varepsilon}, y_i\right) \Pi(dx) + 4\varepsilon. \end{aligned}$$

Also note that

$$\begin{aligned} \bar{\mathbf{E}} \int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 ds &= \sum_{i=1}^{M_\varepsilon} \bar{\mathbf{E}}\left[\mathbf{1}_{\{\bar{X} \in B_i\}} \int_0^T \|\tilde{\psi}_{y_i, \varepsilon}(s, \bar{W})\|^2 ds\right] \\ &= \sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{\mathbf{E}} \int_0^T \|\tilde{\psi}_{y_i, \varepsilon}(s)\|^2 ds \Pi(dx). \end{aligned}$$

Combining above two displays, we have the following estimate on the RHS in (9.4)

$$\begin{aligned} &\bar{\mathbf{E}}\left[\frac{1}{2} \int_0^T \|\bar{\psi}_\varepsilon(s)\|^2 ds + f\left(\bar{W} + \int_0^\cdot \bar{\psi}_\varepsilon(s) ds, \bar{X}\right)\right] \\ &\leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{\mathbf{E}}\left[\frac{1}{2} \int_0^T \|\tilde{\psi}_{y_i, \varepsilon}(s)\|^2 ds + f\left(\tilde{W}^{y_i, \varepsilon}, y_i\right)\right] \Pi(dx) + 4\varepsilon. \end{aligned}$$

It then follows from (9.5), (9.7) and definition of \tilde{K}_ε that the above display can be bounded above by

$$\sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{f}(y_i) \Pi(dx) + 5\varepsilon \leq \sum_{i=1}^{M_\varepsilon} \int_{B_i} \tilde{f}(x) \Pi(dx) + 6\varepsilon \leq \int_{\mathbb{S}} \tilde{f}(x) \Pi(dx) + 7\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we have $\text{RHS} \leq \text{LHS}$ in (9.4) and this completes the proof.

It remains to prove the claim that in proving (9.4) we can assume that $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. The proof of this is similar to arguments in [2, 3] but we provide the details. First note that from [9, Theorem V.16a] we can find $\{f_n\} \subset \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$ such that $\|f_n\|_\infty \leq \|f\|_\infty$ and f_n converges to f a.s. $\bar{\mathbf{P}} \times \Pi$. It then follows from dominated convergence theorem that as $n \rightarrow \infty$, $\tilde{f}_n \rightarrow \tilde{f}$ a.s. Π and hence

$$\int_{\mathbb{S}} \tilde{f}_n(x) \Pi(dx) \rightarrow \int_{\mathbb{S}} \tilde{f}(x) \Pi(dx).$$

To prove the claim, it then remains to show that $\Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f)$ as $n \rightarrow \infty$, where Λ is as in (9.2). Let

$$\mathcal{A}_f(\tilde{\Upsilon}_\Pi) \doteq \left\{ \bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi) : \bar{\mathbf{E}} \int_0^T \|\bar{\psi}(s)\|^2 ds \leq 4\|f\|_\infty \right\}.$$

From the definition of Λ in (9.2) we see that for any bounded f , $-\|f\|_\infty \leq \Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f) \leq \|f\|_\infty$. For any $\bar{\psi}$ such that $\bar{\mathbf{E}} \int_0^T \|\bar{\psi}(s)\|^2 ds > 4\|f\|_\infty$, the right hand side of (9.2) (with g replaced by f) will be strictly larger than $\|f\|_\infty$, which means that such $\bar{\psi}$ does not do better than a ψ in the class $\mathcal{A}_f(\tilde{\Upsilon}_\Pi)$. Thus

$$\Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f_n) = \Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f_n) \text{ and } \Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f) = \Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f).$$

Fix $\varepsilon \in (0, 1)$ and choose $N > 0$ such that $4\|f\|_\infty^2/N \leq \varepsilon/2$. Fix $\bar{\psi} \in \mathcal{A}_f(\tilde{\Upsilon}_\Pi)$ and define the stopping time

$$\tau_N(\bar{\omega}) \doteq \inf \left\{ t \in [0, T] : \int_0^t \|\bar{\psi}(s, \bar{\omega})\|^2 ds \geq N \right\} \wedge T, \quad \bar{\omega} \in \bar{\Omega}.$$

Let $\bar{\psi}_N(s) \doteq \bar{\psi}(s) \mathbf{1}_{[0, \tau_N]}(s)$. Then $\bar{\psi}_N \in \mathcal{A}^N(\tilde{\Upsilon}_\Pi)$, where $\mathcal{A}^N(\tilde{\Upsilon}_\Pi)$ is defined in (9.1), and

$$\bar{\mathbf{P}}(\bar{\psi}_N \neq \bar{\psi}) \leq \bar{\mathbf{P}}(\tau_N < T) \leq \bar{\mathbf{P}} \left(\int_0^T \|\bar{\psi}(s)\|^2 ds \geq N \right) \leq 4\|f\|_\infty/N.$$

Using the above estimate, we have

$$\begin{aligned} \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f_n) &\leq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}_N(s)\|^2 ds + f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}_N(s) ds, \bar{X} \right) \right] \\ &\leq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f_n \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] + \varepsilon. \end{aligned}$$

Taking the infimum over all $\bar{\psi} \in \mathcal{A}_f(\tilde{\Upsilon}_\Pi)$ in the above inequality we have

$$\Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f_n) = \Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f_n) \leq \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f_n) \leq \Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f_n) + \varepsilon.$$

Exactly the same argument with f_n replaced by f gives

$$\Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f) \leq \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f) \leq \Lambda(\mathcal{A}_f(\tilde{\Upsilon}_\Pi), f) + \varepsilon.$$

From Lemma 9.1 we have that as $n \rightarrow \infty$,

$$\Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}^N(\tilde{\Upsilon}_\Pi), f).$$

This proves the claim since $\varepsilon > 0$ is arbitrary. \square

Finally we now give the proof of Proposition 4.1 which extends the result in Lemma 9.2 to a setting with a general filtration.

Proof of Proposition 4.1. Fix $\Pi \in \mathcal{P}(\mathbb{S})$ such that $R(\Pi|\rho) < \infty$. Let $\Upsilon_\Pi \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \bar{P})$ be as in the statement of the proposition and let $\tilde{\Upsilon}_\Pi \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$. It suffices to prove that

$$\begin{aligned} & \inf_{\bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi)} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] \\ &= \inf_{\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \hat{\psi}(s) ds, \bar{X} \right) \right]. \end{aligned} \quad (9.8)$$

We first claim that it suffices to prove (9.8) for $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. The verification of this claim follows along the same lines as the proof of the claim in the proof of Lemma 9.2, and hence we only provide a sketch here. With $\{f_n\} \subset \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$ introduced as below (9.4), it was shown in proof of Lemma 9.2 that $\Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}(\tilde{\Upsilon}_\Pi), f)$. Thus it suffices to show $\Lambda(\mathcal{A}(\Upsilon_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}(\Upsilon_\Pi), f)$ as $n \rightarrow \infty$, where $\Lambda(\mathcal{A}(\Upsilon_\Pi), g)$ is defined through (9.2). For each $N < \infty$, define $\mathcal{A}^N(\Upsilon_\Pi)$ as in (9.1) with $\mathcal{A}(\tilde{\Upsilon}_\Pi)$ replaced by $\mathcal{A}(\Upsilon_\Pi)$. Then by the same stopping time argument below (9.4), it suffices to argue that for each $N < \infty$, $\Lambda(\mathcal{A}^N(\Upsilon_\Pi), f_n) \rightarrow \Lambda(\mathcal{A}^N(\Upsilon_\Pi), f)$ as $n \rightarrow \infty$. However this holds by Remark 9.1. This completes the proof of the claim.

Henceforth we will assume that $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. It is clear that LHS \geq RHS in (9.8). For the reverse inequality, we will show that

$$\begin{aligned} & \inf_{\bar{\psi} \in \mathcal{A}(\tilde{\Upsilon}_\Pi)} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] \\ & \leq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \hat{\psi}(s) ds, \bar{X} \right) \right] \end{aligned} \quad (9.9)$$

for each $\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)$. The proof is similar to that of [3, Lemma 3.5] and we only give a sketch. We will proceed in two steps.

Step 1. Simple $\hat{\psi}$. For simplicity we consider the case where

$$\hat{\psi}(s) = Y \mathbf{1}_{[t, T]}(s),$$

where $t \in [0, T]$, Y is $\hat{\mathcal{F}}_t$ -measurable, and $\|Y\| \leq N < \infty$ a.s. The proof for a general simple process is similar. Consider the map $\varrho : \mathbb{C}([0, t] : \mathbb{R}^d) \times \mathbb{S} \times K_N \rightarrow \mathbb{R}$, where $K_N \doteq \{z \in \mathbb{R}^d : \|z\| \leq N\}$, defined as

$$\varrho(\phi, x, y) \doteq \mathbf{E} \left[\frac{T-t}{2} \|y\|^2 + f \left(\phi^B + \int_0^\cdot y \mathbf{1}_{[t, T]}(s) ds, x \right) \right],$$

where

$$\phi^B(s) \doteq \begin{cases} \phi(s), & s \in [0, t], \\ \phi(t) + B(s-t), & s \in [t, T]. \end{cases}$$

Note that ϱ is bounded, and that by the dominated convergence theorem it is also continuous in (ϕ, x, y) . From a classical measurable selection result (see e.g. [14, Corollary 10.3]) there exists a Borel measurable function $\varrho_1 : \mathbb{C}([0, t] : \mathbb{R}^d) \times \mathbb{S} \rightarrow K_N$ such that

$$\varrho(\phi, x, \varrho_1(\phi, x)) \leq \varrho(\phi, x, y)$$

for all $(\phi, x) \in \mathbb{C}([0, t] : \mathbb{R}^d) \times \mathbb{S}$ and $y \in K_N$. With the definition $\bar{W}_{[0, t]} \doteq \{\bar{W}(s)\}_{0 \leq s \leq t} \in \mathbb{C}([0, t] : \mathbb{R}^d)$ and $\bar{Y} \doteq \varrho_1(\bar{W}_{[0, t]}, \bar{X})$, we set

$$\bar{\psi}(s) \doteq \bar{Y} \mathbf{1}_{[t, T]}(s) \in \mathcal{A}(\bar{\Upsilon}_\Pi).$$

Then

$$\begin{aligned} & \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \hat{\psi}(s) ds, \bar{X} \right) \right] \\ &= \bar{\mathbf{E}} \left\{ \bar{\mathbf{E}} \left[\frac{T-t}{2} \|Y\|^2 + f \left(\bar{W} + \int_0^\cdot Y \mathbf{1}_{[t, T]}(s) ds, \bar{X} \right) \right] \middle| \hat{\mathcal{F}}_t \right\} \\ &= \bar{\mathbf{E}} \varrho(\bar{W}_{[0, t]}, \bar{X}, Y) \\ &\geq \bar{\mathbf{E}} \varrho(\bar{W}_{[0, t]}, \bar{X}, \varrho_1(\bar{W}_{[0, t]}, \bar{X})) \\ &= \bar{\mathbf{E}} \varrho(\bar{W}_{[0, t]}, \bar{X}, \bar{Y}) \\ &= \bar{\mathbf{E}} \left\{ \bar{\mathbf{E}} \left[\frac{T-t}{2} \|\bar{Y}\|^2 + f \left(\bar{W} + \int_0^\cdot \bar{Y} \mathbf{1}_{[t, T]}(s) ds, \bar{X} \right) \right] \middle| \hat{\mathcal{F}}_t \right\} \\ &= \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right]. \end{aligned}$$

So (9.9) holds for simple $\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)$.

Step 2. General $\hat{\psi}$. Next consider an arbitrary $\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)$. We can assume without loss of generality that $\bar{\mathbf{E}} \int_0^T \|\hat{\psi}(s)\|^2 ds < \infty$. Then (cf. [18, Lemma 3.2.4]) there exists a sequence of simple processes $\{\hat{\psi}_n\}_{n \in \mathbb{N}} \subset \mathcal{A}(\Upsilon_\Pi)$ such that

$$\lim_{n \rightarrow \infty} \bar{\mathbf{E}} \int_0^T \|\hat{\psi}_n(s) - \hat{\psi}(s)\|^2 ds = 0.$$

This implies that as $n \rightarrow \infty$,

$$\bar{\mathbf{E}} \sup_{0 \leq t \leq T} \left\| \int_0^t \hat{\psi}_n(s) ds - \int_0^t \hat{\psi}(s) ds \right\|^2 \leq T \bar{\mathbf{E}} \int_0^T \|\hat{\psi}_n(s) - \hat{\psi}(s)\|^2 ds \rightarrow 0.$$

Combining above displays and using Step 1, we have

$$\begin{aligned} & \inf_{\bar{\psi} \in \mathcal{A}(\bar{\Upsilon}_\Pi)} \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\bar{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \bar{\psi}(s) ds, \bar{X} \right) \right] \\ &\leq \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}_n(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \hat{\psi}_n(s) ds, \bar{X} \right) \right] \\ &\rightarrow \bar{\mathbf{E}} \left[\frac{1}{2} \int_0^T \|\hat{\psi}(s)\|^2 ds + f \left(\bar{W} + \int_0^\cdot \hat{\psi}(s) ds, \bar{X} \right) \right], \end{aligned}$$

as $n \rightarrow \infty$, where the convergence holds since $f \in \mathbb{C}_b(\mathcal{C} \times \mathbb{S})$. Thus (9.9) holds for general $\hat{\psi} \in \mathcal{A}(\Upsilon_\Pi)$ and this completes the proof. \square

Appendix A: Proof of Theorem 2.1

Let $\Gamma^n \doteq \frac{1}{n} \sum_{i=1}^n \delta_{(B_i, X_i)}$. Define Q^n and $\tilde{\mu}^n$ as in (5.6) and (5.3) on $\Upsilon = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$, with $\psi_i^n \equiv 0$, $\Pi^n = \theta^{\otimes n}$, and $S_i^n = X_i$. It is then clear that $\mathcal{L}(\Gamma^n, \mu^n) = \mathcal{L}(Q^n_{(1,4)}, \tilde{\mu}^n)$. Note that the bound in (5.5) holds trivially in this case. It follows from Lemmas 5.2–5.5 that $\{(Q^n, \tilde{\mu}^n)\}_{n \in \mathbb{N}}$ is tight in $\mathcal{P}(\Xi) \times \mathcal{D}$, and any weak limit point $(Q^*, \tilde{\mu}^*)$ is almost surely such that $Q^*_{(1,4)} = \mu_0 \otimes \theta$, and $\tilde{\mu}^* = \varpi_{Q^*}$, namely

$$\langle f, \tilde{\mu}^*(t) \rangle = \mathbf{E}^{Q^*} \left[f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \tilde{\mu}^*(s) \rangle ds\}} \right], \quad \forall f \in \mathbb{C}_b(\mathbb{R}^d), t \in [0, T].$$

Recall $\mu(t), a(t), b(t)$ defined in Section 2. From (2.1) $a(t) = \exp \left\{ - \int_0^t a(s)b(s) ds \right\}$. So for $f \in \mathbb{C}_b(\mathbb{R}^d)$ and $t \in [0, T]$,

$$\begin{aligned} \mathbf{E}^{Q^*} \left[f(\tilde{\mathbf{b}}_t) \mathbf{1}_{\{\tilde{\sigma} > \int_0^t \langle \zeta, \mu(s) \rangle ds\}} \right] &= \langle f, \mu_0(t) \rangle \exp \left\{ - \int_0^t \langle \zeta, \mu(s) \rangle ds \right\} \\ &= \langle f, \mu_0(t) \rangle \exp \left\{ - \int_0^t a(s)b(s) ds \right\} \\ &= a(t) \langle f, \mu_0(t) \rangle \\ &= \langle f, \mu(t) \rangle. \end{aligned}$$

So from Lemma 3.1, we must have $\mu = \varpi_{Q^*} = \tilde{\mu}^*$. This implies that $\mu^n \Rightarrow \mu$ as $n \rightarrow \infty$. The result follows. \square

References

- [1] T. Bodineau and M. Lagouge. Large deviations of the empirical currents for a boundary-driven reaction diffusion model. *The Annals of Applied Probability*, 22(6):2282–2319, 2012.
- [2] M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *The Annals of Probability*, 26(4):1641–1659, 1998.
- [3] A. Budhiraja and P. Dupuis. A variational representation for positive functionals of infinite dimensional Brownian motion. *Probability and Mathematical Statistics*, 20(1):39–61, 2000.
- [4] A. Budhiraja, P. Dupuis, and M. Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *The Annals of Probability*, 40(1):74–102, 2012.
- [5] F. Comets, N. Gantert and O. Zeitouni. Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probability theory and related fields*, 118(1):65–114, 2000.
- [6] J. Cvitanović, J. Ma, and J. Zhang. The law of large numbers for self-exciting correlated defaults. *Stochastic Processes and their Applications*, 122(8):2781–2810, 2012.
- [7] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics: An International Journal of Probability and Stochastic Processes*, 20(4):247–308, 1987.
- [8] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications (2nd edition)*, volume 38. Springer Science and Business Media, 2009.

- [9] J. L. Doob. *Measure Theory*. Graduate Texts in Mathematics 143. Springer-Verlag New York, 1994.
- [10] P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*, volume 902 of *Wiley series in probability and mathematical statistics: Probability and statistics*. John Wiley & Sons, New York, 1997.
- [11] Paul Dupuis, Konstantinos Spiliopoulos, and Xiang Zhou. Escaping from an attractor: importance sampling and rest points I. *Ann. Appl. Probab.*, 25(5):2909–2958, 2015.
- [12] Paul Dupuis and Hui Wang. Importance sampling, large deviations, and differential games. *Stoch. Stoch. Rep.*, 76(6):481–508, 2004.
- [13] E. Emrah and C. Janjigian. (2015). Large deviations for some corner growth models with inhomogeneity. *Markov Processes and Related Fields.*, to appear and on arXiv:1509.02234.
- [14] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.
- [15] J. Feng and T. G. Kurtz. *Large Deviations for Stochastic Processes*. Number 131. American Mathematical Soc, 2006.
- [16] P. Freitas. Nonlocal Reaction-diffusion Equations. *Fields Institute Comms*, 21:187–204, 1999.
- [17] G. Jona-Lasinio, C. Landim, and M. E. Vares. Large deviations for a reaction diffusion model. *Probability theory and related fields*, 97(3):339–361, 1993.
- [18] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer New York, 1991.
- [19] T. G. Kurtz. *Approximation of Population Processes*, volume 36 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1981.
- [20] C. Landim and K. Tsunoda. Hydrostatics and dynamical large deviations for a reaction-diffusion model. *arXiv preprint arXiv:1508.07818*, 2015.
- [21] S. T. Rachev. *Probability Metrics and the Stability of Stochastic Models*. Wiley series in probability and mathematical statistics: Applied probability and statistics. John Wiley & Sons, Chichester, 1991.
- [22] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 2006.
- [23] A-S. Sznitman. Topics in propagation of chaos. In *Ecole d’Eté de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Mathematics*, pages 165–251. Springer Berlin Heidelberg, 1991.
- [24] V. Volpert. *Elliptic Partial Differential Equations*. Springer Basel, 2014.

A. BUDHIRAJA

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH

UNIVERSITY OF NORTH CAROLINA

CHAPEL HILL, NC 27599, USA

EMAIL: BUDHIRAJ@EMAIL.UNC.EDU

W-T. FAN

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF WISCONSIN

MADISON, WI 53705, USA

EMAIL: LOUISFAN@MATH.WISC.EDU

R. WU

DIVISION OF APPLIED MATHEMATICS

BROWN UNIVERSITY

PROVIDENCE, RI 02912, USA

EMAIL: RUOYU_WU@BROWN.EDU